

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 1, 2004, 02:44:43 ; Search time 4729.08 Seconds  
(without alignments)  
10669.751 Million cell updates/sec

Title: US-09-913-767-10

Perfect score: 1067

Sequence: 1 agacaagaatggtgaaggct.....gaccataagtgaagttaa 1067

Scoring table: IDENTITY NUC

Gapop 10\_0 , Gapext 1\_0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb\_ba.\*

2: gb\_htg.\*

3: gb\_in.\*

4: gb\_com.\*

5: gb\_ov.\*

6: gb\_pat.\*

7: gb\_ph.\*

8: gb\_pl.\*

9: gb\_pr.\*

10: gb\_ro.\*

11: gb\_sts.\*

12: gb\_sy.\*

13: gb\_un.\*

14: gb\_vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1060.6	99.4	1067	6	AX033553	AX033553 Sequence
2	977.6	91.6	121720	8	AC021044	AC021044 Arabidops
3	476.4	44.6	1049	6	AX033545	AX033545 Sequence
4	447.6	41.9	1075	8	AF078532	AF078532 Arabidops
5	430	40.3	1250	8	AK117664	AK117664 Arabidops
6	428.4	40.1	1102	8	BT005504	BT005504 Arabidops
7	428.4	40.1	1225	6	AX033544	AX033544 Sequence
8	428.4	40.1	1227	8	AF078531	AF078531 Arabidops
9	330	30.9	114144	8	U78721	U78721 Arabidops
10	234.8	22.0	1458	8	AK061667	AK061667 Oryza sat
11	230	21.6	1505	8	AK066610	AK066610 Oryza sat
12	190.6	17.9	1363	8	AK064135	AK064135 Oryza sat
13	182.2	17.1	1476	8	AK071254	AK071254 Oryza sat
14	182.2	17.1	1617	8	AK072660	AK072660 Oryza sat
15	182.2	17.1	175645	8	AP004096	AP004096 Oryza sat
16	150	14.1	110036	8	AC142095	AC142095 Medicago
17	147	13.8	1071	6	AX033550	AX033550 Sequence
18	145.6	13.6	1165	8	AY096558	AY096558 Arabidops
19	145.6	13.6	1428	8	AY074546	AY074546 Arabidops

20	145.6	13.6	3387	6	AX507517	AX507517 Sequence
21	145.6	13.6	3387	6	AX652015	AX652015 Sequence
22	145.6	13.6	82697	8	ATT9A21	AL021713 Arabidops
23	145.6	13.6	194143	8	ATCHRIV48	AL161548 Arabidops
24	141.6	13.3	1152	6	AX653666	AX653666 Sequence
25	141	13.2	14252	8	AF005399	AP005399 Oryza sat
26	136.2	12.8	1194	6	AX033548	AX033548 Sequence
27	135	12.7	1047	6	AX412868	AX412868 Sequence
28	135	12.7	1081	6	AX033549	AX033549 Sequence
29	135	12.7	1106	8	AF370622	AF370622 Arabidops
30	133.6	12.5	73179	8	AP006383	AP006383 Lotus cor
31	133	12.5	14894	8	AC121489	AC121489 Oryza sat
32	133	12.5	195186	2	AC149309	AC149309 Zea mays
33	131.4	12.3	1200	6	AX654080	AX654080 Sequence
34	128.6	12.1	1506	8	AK105352	AK105352 Oryza sat
35	127	11.9	96046	8	AC005967	AC005967 Arabidops
36	119.8	11.2	1293	6	AX033547	AX033547 Sequence
37	119.8	11.2	58427	8	ATF15J5	AL110123 Arabidops
38	116.6	10.9	1145	6	AX033546	AX033546 Sequence
39	114	10.7	1117	8	AY096468	AY096468 Arabidops
40	114	10.7	1333	8	AY074535	AY074535 Arabidops
41	114	10.7	83698	6	AX540653	AX540653 Sequence
42	114	10.7	83698	8	AB010072	AB010072 Arabidops
43	114	10.7	104386	8	ATT32A17	AL161813 Arabidops
44	114	10.7	193362	8	ATCHRIV24	AL161512 Arabidops
45	112.8	10.6	1251	8	BT003358	BT003358 Arabidops

ALIGNMENTS

RESULT 1	AX033553	1067 bp	DNA	linear	PAT 21-SEP-2000
LOCUS	AX033553	Sequence 10 from Patent WO0049152.			
DEFINITION	AX033553				
ACCESSION	AX033553.1	GI:10280295			
VERSION	AX033553.1				
KEYWORDS	Arabidopsis thaliana (thale cress)				
SOURCE	Arabidopsis thaliana				
ORGANISM	Arabidopsis thaliana				

Arabidopsis thaliana (thale cress)  
Eukaryota, Viridiplantae, Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi

REFERENCE 1  
AUTHORS Andre B., Buerkle, L., Frommer, W.B. and Gillissen, B.  
TITLE Nucleic acids that code for a nucleobase transporter  
JOURNAL Patent: WO 0049152-A 10 24-AUG-2000;  
ANDRE BRUNO (BE) ; BUERKLE LUKAS (DE) ; FROMMER WOLF B (DE) ;  
GILLISSEN BERND (DE)

FEATURES source  
1..1067  
Location/Qualifiers  
/organism="Arabidopsis thaliana"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:3702"

ORIGIN  
Query Match 99.4%; Score 1060.6; DB 6; Length 1067;  
Best Local Similarity 99.6%; Pred. No. 5.2e-269;  
Matches 1063; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy	1	AGACAAGAATGGTGAAGGCTCTTGTGATCATATAAATGCATTAATTCACCATAGGAAACT	60
Db	1	AGACAAGAATGGTGAAGGCTCTTGTGATCATATAAATGCATTAATTCACCATAGGAAACT	60
Qy	61	GTGAGGTCCTTTGATATGCTCTACTTCAACAATGGCGGTAAGGATTTGGTTCT	120
Db	61	GTGAGGTCCTTTGATATGCTCTACTTCAACAATGGCGGTAAGGATTTGGTTCT	120
Qy	121	CTAGCTTTCTTGAACACTGCAGGCTTTCCTGTTATCTTCATTCCTCTCTCTTCTTACA	180
Db	121	CTAGCTTTCTTGAACACTGCAGGCTTTCCTGTTATCTTCATTCCTCTCTCTTCTTACA	180
Qy	181	TTACCCGGCGGAAGCAACAATGTGGGTGATAGTACAAAGTTCTTCTTCTTATCAAAACCGC	240

Db	191	TTACCGCGCGGAGAACCAATGTTGGGTGATAGTACAAAGTTCTTCTTATCAACCGC	240
Qy	241	GTCTTCTTATCGCGCTGTTATTTAGGACATCTCTCAGGGTTTGATAACTACTTGTATG	300
Db	241	GTCTTCTTATCGCGCTGTTATTTAGGACATCTCTCAGGGTTTGATAACTACTTGTATG	300
Qy	301	CATATGGTATAGCTTATCTTCAGTTTCTACAGTCTCTTATCATATGTTCTTCAGTTAG	360
Db	301	CATATGGTATAGCTTATCTTCAGTTTCTACAGTCTCTTATCATATGTTCTTCAGTTAG	360
Qy	361	CTTTTATAGCTATCTTCTCATCTTTCATGTTTAAACATAAGTTCACCTCTTTTACCATCA	420
Db	361	CTTTTATAGCTATCTTCTCATCTTTCATGTTTAAACATAAGTTCACCTCTTTTACCATCA	420
Qy	421	ATGCTGTTGTGTGTGACTGTGTGCTGCGGTTTGGGAATGCATACCGAACTGATA	480
Db	421	ATGCTGTTGTGTGTGACTGTGTGCTGCGGTTTGGGAATGCATACCGAACTGATA	480
Qy	481	AGCCAGTTTCATGAGCTCACAGCAGTACATAACTGTTCTTGATTTACTGTACGAGCAG	540
Db	481	AGCCAGTTTCATGAGCTCACAGCAGTACATAACTGTTCTTGATTTACTGTACGAGCAG	540
Qy	541	CTGTTATGATGCTTTTCATCTTCCATTTAGTGGAACTTGCTTACAGAAAGCTTAAGCAA	600
Db	541	CTGTTATGATGCTTTTCATCTTCCATTTAGTGGAACTTGCTTACAGAAAGCTTAAGCAA	600
Qy	601	CCATGAGCTATACCCCTTGCTCGAGTTCCAGTTGATTTTGTTGCTCTCTTCTATTG	660
Db	601	CCATGAGCTATACCCCTTGCTCGAGTTCCAGTTGATTTTGTTGCTCTCTTCTATTG	660
Qy	661	TCAGGCTCATCGTATGTTTCATCGCTGCTGATTTTCAAGCAGCCCTTACCAAGAGCAA	720
Db	661	TCAGGCTCATCGTATGTTTCATCGCTGCTGATTTTCAAGCAGCCCTTACCAAGAGCAA	720
Qy	721	GAGAGTTTCAAGCTTGGAGAGCATTTGTTATGTTGGTGGCTGTTTTCAGCCATCATAT	780
Db	721	GAGAGTTTCAAGCTTGGAGAGCATTTGTTATGTTGGTGGCTGTTTTCAGCCATCATAT	780
Qy	781	GGCAAGGCTTCTTCTTGGAGGCCATTGATTAATCTTCCACATCGTCTCTCGTCGG	840
Db	781	GGCAAGGCTTCTTCTTGGAGGCCATTGATTAATCTTCCACATCGTCTCTCGTCGG	840
Qy	841	GTATATCATATCAGTGTGTTTGGCAATTACAGAGTTTTCAGCTGTATATCTTACCATG	900
Db	841	GTATATCATATCAGTGTGTTTGGCAATTACAGAGTTTTCAGCTGTATATCTTACCATG	900
Qy	901	AAAAGTTTCAAGCTGAGAGGCACTTCTCTTGCTCTCTCCCTTTGGGGCTTTGCTCTT	960
Db	901	AAAAGTTTCAAGCTGAGAGGCACTTCTCTTGCTCTCTCCCTTTGGGGCTTTGCTCTT	960
Qy	961	ACTTTTATGTTGAGATAAAGTCTGGCAGGATAAAGGAGAATTCAGCAGAGGAGATC	1020
Db	961	ACTTTTATGTTGAGATAAAGTCTGGCAGGATAAAGGAGAATTCAGCAGAGGAGATC	1020
Qy	1021	AGGAGACAGACAATCTTCTTGTCTCAAGACCCATAAGTGAGTGTAA	1067
Db	1021	AGGAGACAGACAATCTTCTTGTCTCAAGACCCATAAGTGAGTGTAA	1067
RESULT 2			
AC021044			
LOCUS	AC021044	121720 bp	DNA linear PLN 30-OCT-2002
DEFINITION	Arabidopsis thaliana chromosome I BAC F3H9 genomic sequence,		
	complete sequence.		
ACCESSION	AC021044		
VERSION	AC021044.5	GI:8347959	
KEYWORDS	HTG.		
SOURCE	Arabidopsis thaliana (thale cress)		
ORGANISM	Arabidopsis thaliana		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;		
	rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.		

REFERENCE	1	(bases 1 to 121720)	Federspiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F., Altafi,H., Nguyen,M., Lam,B., Southwick,A., Miranda,M., Brooks,S., Buehler,E., Chao,Q., Chin,C., Chlou,J., Choi,E., Dunn,P., Gonzalez,A., Howng,B., Kim,C., Koo,T., Lee,J.M., Lenz,C., Khan,S., Kim,C., Koo,T., Lee,J.M., Lenz,C., Liu,A., Liu,S., Mukharsky,N., Pham,P., Sakano,H., Shinn,P., Toriumi,M., Vaysberg,M., Yu,G., Ecker,J., Theologis,A. and Davis,R.W.
JOURNAL			Unpublished
REFERENCE	2	(bases 1 to 121720)	Federspiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F., Altafi,H., Nguyen,M., Lam,B., Southwick,A., Bei,Q., Buehler,E., Chin,C., Chlou,J., Choi,E., Dunn,P., Gonzalez,A., Howng,B., Kim,C., Koo,T., Lee,J.M., Lenz,C., Li,J., Liu,A., Liu,K., Liu,S., Mukharsky,N., Pham,P., Sakano,H., Schwartz,J., Shinn,P., Thavari,A., Toriumi,M., Vaysberg,M., Walker,M., Yu,G., Ecker,J., Theologis,A. and Davis,R.W.
JOURNAL			Direct Submission
TITLE			Submitted (12-JAN-2000) DNA Sequencing and Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA
REFERENCE	3	(bases 1 to 121720)	Federspiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F., Altafi,H., Nguyen,M., Lam,B., Southwick,A., Bei,Q., Buehler,E., Chin,C., Chlou,J., Choi,E., Dunn,P., Gonzalez,A., Howng,B., Kim,C., Koo,T., Lee,J.M., Lenz,C., Li,J., Liu,A., Liu,K., Liu,S., Mukharsky,N., Pham,P., Sakano,H., Schwartz,J., Shinn,P., Thavari,A., Toriumi,M., Vaysberg,M., Walker,M., Yu,G., Ecker,J., Theologis,A. and Davis,R.W.
JOURNAL			Direct Submission
TITLE			Submitted (29-MAR-2000) DNA Sequencing and Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA
REFERENCE	4	(bases 1 to 121720)	Federspiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F., Altafi,H., Nguyen,M., Lam,B., Southwick,A., Bei,Q., Buehler,E., Chin,C., Chlou,J., Choi,E., Dunn,P., Gonzalez,A., Howng,B., Kim,C., Koo,T., Lee,J.M., Lenz,C., Li,J., Liu,A., Liu,K., Liu,S., Mukharsky,N., Pham,P., Sakano,H., Schwartz,J., Shinn,P., Thavari,A., Toriumi,M., Vaysberg,M., Walker,M., Yu,G., Ecker,J., Theologis,A. and Davis,R.W.
JOURNAL			Direct Submission
TITLE			Submitted (08-JUN-2000) DNA Sequencing and Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA
REFERENCE	5	(bases 1 to 121720)	Federspiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F., Altafi,H., Nguyen,M., Lam,B., Southwick,A., Ecker,J., Theologis,A. and Davis,R.W.
JOURNAL			Direct Submission
TITLE			Submitted (11-AUG-2000) DNA Sequencing and Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA
REFERENCE	6	(bases 1 to 121720)	Federspiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F., Altafi,H., Nguyen,M., Lam,B., Southwick,A., Ecker,J., Theologis,A. and Davis,R.W.
JOURNAL			Direct Submission
TITLE			Submitted (15-AUG-2000) DNA Sequencing and Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA
COMMENT			On Jun 8, 2000 this sequence version replaced gi:7340332. Bases 1-7,373 of clone F3H9 overlap with bases 96,792-104,163 of IGF BAC clone F3M18, gb AC010155 e-mail for correspondence: arab@sequence.stanford.edu Genes with similarity to proteins in the databases are named 'putative', '-like' or 'similar to'. Genes that have EST similarity but no significant protein similarity are described as 'unknown proteins'. Genes that are annotated based only on gene prediction software are described as 'hypothetical proteins'. The gene prediction programs used to predict genes include: Grail (informatics Group, Oak Ridge National Laboratory, <a href="http://compbio.ornl.gov/section/index.html">http://compbio.ornl.gov/section/index.html</a> ), GENSCAN (Chris Burge,



```
http://gnomic.stanford.edu/~chris/GENSCANW.html), and NetPlantGene
(S.M. Hebsgaard, et al., CBS, Technical University of
Denmark, http://www.cbs.dtu.dk/NetPlantGene.html).
FEATURES
    Location/Qualifiers
        1..121720
            /organism="Arabidopsis thaliana"
            /mol_type="genomic DNA"
            /db_xref="taxon:3702"
            /chromosome="I"
            /clone="F3H9"
        misc_feature
            1..7373
                /note="overlap with bases 96,792-104,163 of IGF BAC clone
                F3M18, gb|AC010155, see GenBank record for clone F3M18 for
                annotation in this region."
        gene
            9852..10816
                /gene="F3H9.1"
                /note="synonym: DRM1"
        CDS
            join(9852..9957,10049..10274,10780..10816)
                /genes="F3H9.1"
                /note="Location of cDNA clone gb|AF053746 and genomic
                clone gb|AF053747."
                /codon_start=1
                /product="dormancy-associated protein"
                /protein_id="AAF98422.1"
                /db_xref="GI:9795604"
            /translation="WVLEKLMWDVVAGPQDRGLRLKITTQPINIRDIGEGSSSK
            VMHSLRTPAAVSPGTPPTPTTPTRKDNVWRSVFNPGSNLATRAIGSNIFDKPTH
            PNSPSVYDMLYSGDSRSQHR"
            complement(12229..14900)
        gene
            /gene="F3H9.2"
        CDS
            complement(join(12229..12924,13048..13089,13397..13665,
            14024..14447,14511..14900))
                /gene="F3H9.2"
                /note="Hypothetical protein"
                /codon_start=1
                /protein_id="AAF98443.1"
                /db_xref="GI:9795625"
            /translation="MDSRDMLACEDIEYSIKKNETRSLVDLPFGIKPIGLKWWFKLK
            RNSDYSINKYKARLVAGQVQRYGDFVEVPVACIEKIIIVDLSEANDWEIOIDVK
            TAPLHGELDTVTTPEDFVEKGEGLNQIQLQLOFEKCSKSPVYKRTISGEFLL
            VAVYVDLFTCTNTIIDEFKGMTSNFMSNLKLYTLYGLIEVPQKEEIRLNQSR
            YEMKILEETAMVCSLIHTSRESVQLAKSENEKIDATFRKVKGCPLRYLLNLSE
            VTGLSPVRVIRIDKNSALSLTNPFVFGHEGKHRRHYFRERVENQIIEVHVIGI
            EQRADILTKALVRMFKEMRDLIGVQDELISQEIINIQQOASMYOTLKRQATDS
            GKRMSYFVRIRDPDTCQCI FRPIPFKGMKVRKHVDDEGSGGSKVKVRLTSVSG
            DGLSLSPENETTTIRNEKRLMLEIFQPESTDSDSDESGTEEFEMETNLSLVSLOW
            GKGKRSMEKGTWKHLLHNHVQKHNTQDVEVKNCLYCEITLVHSFSLIGVSMQMOMD
            IOVLVSMDSAVIPNLQNPNQVSKCPNVLTKSLCEVCLILT"
            20721..24395
                /gene="F3H9.3"
        CDS
            join(20721..20817,21062..21222,21318..21365,21455..21514,
            21709..21837,21988..22098,22183..22286,22430..22488,
            22580..22674,22757..23426,23624..23816,23957..24048,
            24148..24395)
                /gene="F3H9.3"
                /note="Unknown protein"
                /codon_start=1
                /protein_id="AAF98423.1"
                /db_xref="GI:9795605"
            /translation="MDVSKVVSFRNFAVLVRKEGVPSPDLSPNSSSGILLPRDIFL
            SGEVAAKLVFEAGQDMALVLTVAVVEPFLTIGHRTSSISQDPVKLIIPGAMIEIMVE
            GQLKSEKAPFWPAPQLSLVDVPVSAALQSLIEASSGSKSDGWDIGWSLVSAANGS
            QPSINIEHYPKLMQDLDEPHANFMAKSAATMAILGVPLSLILGQPSMNFASSSKGD
            LVALGSPGLISPVNFNSYPSSGLKSLIMADVRLPQMEGAPVFANKHILIGLIR
            PLRKNSGVEIQQLVVPWCAITATCSHLLLEEPSVEGKASQWSVLSVSKSDASIPAQ
            VAIEKAMESVCLITVNDGVASGIIINHEGLILTNHLLPEWRYGKGVYGEYGFYV
            LGAEFFSGVFNQKQSLTPKAPRNHYSSVGENIREYKHNFLTQTHDIRVRLCH
            LDSWTVCANVYVIKEQDIALQLLEVVPGLQPIITANFSSPLGTTGHRVHVLGFG
            PCGLSPSICSGWAKVHAKRLNTQSIQSEVAFPAMLETTAAVHPGSGGAVLNS
            SCHMIGLVTNSAHGAGTVIPLNFSIPCAVLPIFKAEQDMQNTIILQTLQDSEEL
            SSIALMPSLSFKTEQSLPNLPKLKDGNNKQTKGSQFAKFIATQDMFVFRPKLUSD
            VIPSKU"
            31599..32534
```

```

/gene="F3H9.4"
31599..32534
/gene="F3H9.4"
/note="Hypothetical protein"
/codon_start=1
/protein_id="AAF98424.1"
/db_xref="GI:9795606"
/translation="MQSKNMIVASSHQOQOQPOQPOQLKPCPCDSNNTKFCYNN
YSLQPHFCACKRYTRGGTLLRNVPVGGSVYKRNKVRPSTATTATTTASTVSTNSS
SPNNPHQISHPSSNMNHPFLFYGLSDHMSCCNNLPMIPSRPSDSKTKCSSGLESEFL
SSGSPSSALGGLGLPHQMSHDHTINGSFINNTNNKPELLSGLFGSSSSSSSTLQHP
HKPNNGDMIGQSHLQTLASLDLHVGNNEDMKYKREGKLDQISGNINGFMSSSSSL
DPSYNNNMNNAVVVNGAWLDPTNNNVGSSLSLI"
45230..48424
/gene="F3H9.5"
join(45230..45724,45882..45956,46851..46951,47136..47182,
47292..47365,47715..47984,48407..48424)
/gene="F3H9.5"
/note="Hypothetical protein"
/codon_start=1
/protein_id="AAF98425.1"
/db_xref="GI:9795607"
/translation="MDNPLPFPSSNANSVQELSHMDPNNRNHSHTTVPYDHHQAPHH
FLPFPSPYEQMAAAMNPQPVYLSECYPOIPIVOTGTSEFSLVGNPCLVGGGFLDP
RMTKARINRNKAMMRNRNNSPNSPSELVDSKRLMLNKNVQISDKKDSYQOS
TFDNKLRVLCEKELKNSDVGSLGRIVLPRKDAEANLPKLSDEKGIYVQMRDVFMSQS
WSFKYKFWNSKNRWYVLENTGEFVKQNGAIGDPLTIYEDESKNLYFAMNGNSKON
EGRENEGRNHYBEAMLDYIIPRDEBEASIAMLGNLNDHYPIPNDLMDLTDLQHHQ
ATSSMPEDHAYVGSDDQVLDML"
54225..55737
/gene="F3H9.6"
join(54225..54993,55427..55737)
/gene="F3H9.6"
/note="Unknown protein; proline rich amino acid sequence"
/codon_start=1
/protein_id="AAF98426.1"
/db_xref="GI:9795608"
/translation="MGFIGKSLVSLVALWCFTSPVSTEEVNHKTQTOTPSLAPAPAYH
HGHHHPHPHHHPHPHPHPHPKAPSPVKPVKAPSPVPKPPVPPVPTKAPKVPK
PTKPPVPKPSVPAKPPVKPVPTKAPKVPKPPKPPKVPKPPKPPKPPKPPKPPV
KPPVPTKAPKVPKPPKPPKPPKPPKPPKPPKPPKPPKPPKPPKPPKPPKPPKPP
PVYPPKRNLSLVAVRGTVYCKSKYAAPFTLLGKAPIEGATVKKPVPSKPPITATBT
DKNGYFLLAPKTVTFNFGRCRVLYKSKDYKCSKVSCLFGGVDVGAELKPEKLGKS
Query Match          91.6%; Score 977.6; DB 8; Length 121720;
Best Local Similarity 93.2%; Pred. No. 4.8e-247;
Matches 1063; Conservative 0; Mismatches 4; Indels 73; Gaps 1;
QY 1 AGACAAGAATCGTGAAGGCTCTTGTGATCATAAACTGCATAATTCTAGCCATAGGAACT 60
Db 83696 AGACAAGAATCGTGAAGGCTCTTGTGATCATAAACTGCATAATTCTAGCCATAGGAACT 83695
QY 61 GTGGAGGTCTTTGATNATGCGTCTTACTTCAACAATAGCGCGTAAAGGATTTGGTTCT 120
Db 83686 GTGGAGGTCTTTGATNATGCGTCTTACTTCAACAATAGCGCGTAAAGGATTTGGTTCT 83745
QY 121 CTAGTCTTTGAAACTGCAGGCTTTCCTGTTATCTTCAATTCCTCTCTCTCTTACA 180
Db 83746 CTAGTCTTTGAAACTGCAGGCTTTCCTGTTATCTTCAATTCCTCTCTCTCTTACA 83805
QY 181 TTACCCGCGCAGAGCAACAATAGTGGGTGATAGTACAAAGTTCTTTCTTATCAACCGC 240
Db 83806 TTACCCGCGCAGAGCAACAATAGTGGGTGATAGTACAAAGTTCTTTCTTATCAACCGC 83865
QY 241 GTCTCTTATCGCGCTGTTATTGTTAGGCATTCTCTCAGGGTTTGATTAACCTTGTATG 300
Db 83866 GTCTCTTATCGCGCTGTTATTGTTAGGCATTCTCTCAGGGTTTGATTAACCTTGTATG 83925
QY 301 CATATGGTATAGCTTATCTCCAGTTTCTACAGCTGCTCTTATCATCTGCTTCCACAGTTAG 360
Db 83926 CATATGGTATAGCTTATCTCCAGTTTCTACAGCTGCTCTTATCATCTGCTTCCACAGTTAG 83985
QY 361 CTTTATAGCTATCTTCTCTCATTTCTTTCATGGTTAAACATAAAGTTCACTCTCTTTTACCATCA 420
```





Kawai, J., Hayashizaki, Y. and Shinozaki, K.  
Arabidopsis thaliana full-length cDNA  
Published Only in Database (2002)  
22 (bases 1 to 1250)  
Seki, M., Iida, K., Satou, M., Sakurai, T., Akiyama, K., Ishida, J.,  
Nakajima, M., Enju, A., Kamiya, A., Narusaka, M., Carninci, P.,  
Kawai, J., Hayashizaki, Y. and Shinozaki, K.  
Direct Submission  
Submitted (25-Nov-2002) Motoaki Seki, RIKEN Genomic Sciences  
Center; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa  
230-0045, Japan [E-mail: mseki@gsc.riken.go.jp,  
http://pfweb.gsc.riken.go.jp, Tel: 81-45-503-9625,  
Fax: 81-45-503-9586]  
An Arabidopsis full-length cDNA library was constructed essentially  
as reported previously [Seki et al. (1998) Plant J. 15:707-720];  
Seki et al. (2002) Science 296:141-145]. cDNA cleaved with BamHI  
and XhoI was ligated to modified lambda FLC-1-E vector [Carninci et  
al. (2001) Genomics 77:79-90] digested with BamHI and SalI.  
This clone is in a modified pBluescript vector.  
Please visit our web site (<http://pfweb.gsc.riken.go.jp/>) for  
further details.

FEATURES	source
Location/Qualifiers	
1. .1250	
/organism="Arabidopsis thaliana"	
/mol_type="mRNA"	
/db_xref="taxon:3702"	
/chromosome="1"	
/clone="RAF17-32-O17"	
/scotype="Columbia"	
/note="common name: thale cress"	
1. .1250	
/gene="At1g28230/F3H9_10"	
53. .1123	
/gene="At1g28230/F3H9_10"	
/codon_start=1	
/product="putative purine permease"	
/protein_id="BAC42317.1"	
/db_xref="GI:26450405"	
/translation="MKNGLIILIIICITLTGCGPLLTRLVFTNGKRIWFMSFLSTA GPILLIPLVLSFRRSRNPNNAENKRTKFLMETPLFIASVITGLLGLDNL YSGLAYLPVSTSLIQLGFAPNALFALLVKQKTFPSINAVITLVGIGLIALHS DGDPAKESKBYVGFMTVAALYAFILPLVETTKARQEIITFPLVLEIQMVC LAATFCVGMITVGFQVIAREAREFKIGSVFYVALIVITGIIWQGFFLGAIVF CASSLASGLVILLVLPVTEVFAVVCPRKFNQAEKGVSLLSLWGFVSYFYGEPKSKG VYDKPOPEETELPLPVSQVYA"	

ORIGIN	Query Match	Score	DB	Length
	Best Local Similarity	66.5%	Pred. No. 1.7e-102;	
	Matches 673; Conservative	0; Mismatches 315; Indels	24; Gaps	3;
QY	1	AGACAAGAAATGGTGAAGGCTCTTGTGATCATAAACTGCATAATCTTAGCCATAGGAAACT	60	
DB	45	AGAGAAGATGAAGATGGTTTGATTAATCATAAACTGTATTATCCTCACTATAGGAACAT	104	
QY	61	GTGGAGGTCCTTTTGATTAATGCGTCTCTACTTCAACAATGCGCGGTAAAGAAATTTGTTCT	120	
DB	105	GTGGAGGTCCTTTGTTAACTCGTCTCTACTTCAACAATGCGCGAAACGAATCTGGTTCA	164	
QY	121	CTACGTTTCTGAAACTGAGGCTTTCCTGTTATCTTCAATCTCTCTGCTCTTCTCTTACA	180	
DB	165	TGAGCTTCTCTATCAACGCGTGGTTTTCCAATCATCTCATCCCTCTCTTGGTCTCCTTCC	224	
QY	181	TTATCCGGCGCAGAACCAACAATGTGGTGA-----TAGTACAAAGTT	222	
DB	225	TACGCCGTGCGCGACAGAACCGCAACCTTAACAACGCGGAAACAAGCGGAAACAAAGC	284	
QY	223	TCTTTCTTATCAAAACGCGCTCTTCTTATGCGCGCTGTATTGTAGGCAATCTCTCAGGGT	282	
DB	285	TCTTCTCATGGAAACTCCTCTGTTTATCGCTTCCATTGTTCATAGGGTTGCTCACAGGAC	344	
QY	283	TTGATAACTACTTGTTATGCATATGGTATAGCTTATCTTCCAGTTTCTACAGCTGCTCTTA	342	

345	TTGACAACTACTTATATCTTTACGGATTAGCATATCTGCCAGTTTCAACTTTCATCGCTCA	404
343	TCATTGCTTCAACAGTTAGCTTTTATAGCTATCTTCTCATTTCTTCATGGTTAAACAATAAGT	402
405	TAATCGAACTCAACTAGCTTTCAACGCTCTCTTCGCTTCTTGTGTAGTCAAGCAAAAGT	464
403	TCATCTCTTTTACATCAATAGCTGTGTGTGTGTGTGACTGTGTGGTGTGCGGTTTTGGAA	462
465	TCATCTCGTCTCTCCATAAAGCCGCTCGTTTTGTGTGACGGTTGGTATCCGGATCCCTGCGT	524
463	TGCATATCCGAAACTGATAAAGCCAGTTTCATCAGACTCAACAGCAGTACATAACTCGTTTCT	522
525	TACACAGTGATGAGACAAACCGGCTAAGAGAGCAAGAAAGATATGTGGTGGGTTCT	584
523	TGATTAAGTGTGAGCAGCTGTTATGATGCTTTTCATCTTTGCCATTAAGTGGAACTTCGTT	582
585	TGATGACTGTGGTTGACGCTCTTCTCTATGCTTTTATATACCGCTCGTTGAGCTAACTT	644
583	ACCAGAAAGCTAAGCAAAACCATGAGCTATACCTGTGCTCGAGTTCGAGTTGATTTTGT	642
645	ACAAGAAAGCTCGTCAAGAAATCACTTCCACTGTGTGTGAGATTCAGATGTGCATGT	704
643	GTCTCTCTGCTTCTATTGTCCAGCGTCATCGGTATGTTTCATCGCTGGTGTGATTTCAAGCAGG	702
705	GCCTTGTGCTACTTTTTTCTGTGTCAATGGCATGTTTCATCGTTGGAGATTTTA--AGG	761
703	CCTTACCAAAAGAACGAGAGCTTCAAGCTTGGAGAGGCATTTGTT---CTATGTGGTGG	759
762	TGATAGCAAGAGAGCAAGAGAGTTCAAGATTCGGAGATCAGTGTTTTACTATGCAATTGA	821
760	CTGTGTTTTCAGGCCATCATATGGCAAGGCTTCTTTCTGGAGCCATTCGGATTAATCTTCT	819
822	TAGTGATCAGAGAAATAATGGCAAGGTTTCTTTAGGAGCCATAGGGATTCGTGTTTT	881
820	CCACATCGTCTCTCGTCTCGGTTATTATGATATCAGTGTCTTTTGGCCAAATTCAGAGGTTT	879
882	GTGCAATCATCACTAGCTTCTGGTGTTCGTATAAGTGTTCGTCTCCGGTGAAGGATTT	941
880	TAGCTGTATATTTCTACCATGAAAGTTTCAAGCTGAGAGGGACTTTCTCTTCTGCTCTCT	939
942	TCGCGCTGTTTGTTTCCGGGAGAGTTTCAGCAGAGAAAGTGTCTCTCTACTTCTTT	1001
940	CCCTTTGGGCTTTGTCTCTTTACTTTTTATGGTGAGATAAAGTCTGGCGAGGA	991
1002	CTCTTTGGGGATTGTCTCTTACTTCTACGGCGAGTTTAAATCCGGCAAGAA	1053

RESULT	6
LOCUS	BT005504
DEFINITION	Arabidopsis thaliana clone U50269 putative purine permease (Atlg28230) mRNA, complete cds.
VERSION	BT005504.1
KEYWORDS	FLI CDNA.
SOURCE	Arabidopsis thaliana (thale cress)
ORGANISM	Arabidopsis thaliana
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliopsida; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis. 1 (bases 1 to 1102) Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Hsuan,V.W., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C., Toriumi,M., Wong,C., Wu,H.C., Yu,G., Yuan,S., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Kawai,J., Kim,C.J., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A., Tripp,M.G., Wu,T., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A. Arabidopsis Open Reading Frame (ORF) Clones Unpublished 2 (bases 1 to 1102) Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Hsuan,V.W., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Wong,C., Wu,H.C., AUTHORS
JOURNAL	TITLE
REFERENCE	2 (bases 1 to 1102)
AUTHORS	Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Hsuan,V.W., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Wong,C., Wu,H.C.,



ANDRE BRUNO (BE) ; BUERKLE LUKAS (DE) ; FROMMER WOLF B (DE) ;  
GILLISSEN BERND (DE)  
FEATURES  
source  
location/Qualifiers  
1..1225  
/organism="Arabidopsis thaliana"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:3702"

ORIGIN

Query Match 40.1%; Score 428.4; DB 6; Length 1225;  
Best Local Similarity 66.4%; Pred. No. 4.6e-102;  
Matches 672; Conservative 0; Mismatches 316; Indels 24; Gaps 3;  
Qy 1 AGACAAGATGGTGAGGCTCTTGATCATATAACTGCATAATCTACCCATAGAAACT 60  
Db 17 AGAAGAATGAAGAATGGTTGATAATCATATAAATCTGTAATCTCTACTATAGAACAT 76  
Qy 61 GTGGAGGTCTTTGATTATGCGTCTCTACTTCAACAATGGCGGTAAGAGGATTTGGTTCT 120  
Db 77 GTGGAGGTCTTTGTTAACTGCTCTACTTCAACCAATGGCGGAACGAATCTGGTTCA 136  
Qy 121 CTACGTTTCTTGAACACTCAGGCTTTCCTGTATCTTCAATTCCTCTCTCTCTTCTTACA 180  
Db 137 TGAGCTTCTATCAACCGCTGGTTTCCAAATCATCTCATCCCTCTCTTGGTCTCCTTCC 196  
Qy 181 TTACCGGGGAGACCAATATGGGTGA-----TAGTACAAGTT 222  
Db 197 TCAGCGGTGGCGGCAACCGCAACCTTAACAACGCGGAAACAAGCGGAAACAAGC 256  
Qy 223 TCTTTCTTATCAACCGGCTCTCTTATCGCGGCTGTTATGTAGCGATTCTCTCAGGGT 282  
Db 257 TCTTCTCATGGAACCTCTCTGTTATCGCTCCATGTCTAGGGTTGCTCACAGGAC 316  
Qy 283 TTGATAACTACTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 342  
Db 317 TTGACAACTACTTATCTTACGGATTAGCATATCTGCCAGTTTCAACTTTCATCGCTCA 376  
Qy 343 TCATTGCTTCACAGTTAGCTTTATAGCTATCTTCTCATCTTCTCATGTTTAACTAAGT 402  
Db 377 TAATCGGAACCAACTAGCTTTCACAGCTCTCTCTCGCTCTTCTGTTAGTCAAGCAAAAGT 436  
Qy 403 TCACCTCTTTTACCATCAATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 462  
Db 437 TCACCTCGGTTCCCATAAACCGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 496  
Qy 463 TGCATACCGAACTGATAGGCAAGTTTCATGAGACTCAACAAGAGTACATAACTGGTTTCT 522  
Db 497 TACACAGTCAATGGAGCAAAACCGGCTAAGGAGAGCAAGAAAGAGTATGTGGTGGTTCT 556  
Qy 523 TGATTACTGTAGCAGCAGCTGTATGATGATGATGATGATGATGATGATGATGATGATGAT 582  
Db 557 TGATGACTGTGGTTGCAAGCTCTCTCTATGCTTTTATATATACCGCTCGTTCAGCTAACTT 616  
Qy 583 ACCGAAAGCTAAGCAAAACCACTAGCTATACCTTGTCTCGAGTTCAGATTGATTTTGT 642  
Db 617 ACAAGAAAGCTGTCAGAAATCACTTCCCACTGTGCTTGAGATTGAGATGCTCATGT 676  
Qy 643 GTCTCCTTGTCTTATGTCAGCGTCAATGATGATGATGATGATGATGATGATGATGATGATG 702  
Db 677 GCGTTGCTGCTACTTTTCTGTGTCATGGCATGTTTCATGCTGGAGATTTTA---AGG 733  
Qy 703 CCTTACCAAAAGACAGAGAGTTTCAAGCTTGGAGAGCATGTT---CTATGTGGTG 759  
Db 734 TGATAGCAAGAGACAGAGAGTTTCAAGATGGAGGATCAAGTTTGTATGATGATGATGATG 793  
Qy 760 CTGTGTTTTTCAGCCATCATATGGCAAGGCTTCTTCTGGGAGCCATTCGATTAATCTTCT 819  
Db 794 TAGTATCACAGGAATATATGGCAAGGTTTCTTCTTAGGAGCCATAGGATTTGTTTT 853  
Qy 820 CCACATCGTCTCTCGGTATATATGATATCATAGTCTTTTGGCAATTCAGAGGTTT 879  
Db 854 GTGCATCATCACTAGCTTCTGGTGTCTGATAAGTGTCTGCTTCGCGTCACTGAAAGTTT 913

Qy 880 TAGCTGTTATATTCTACCATGAAAAGTTTCAAGCTGAGAAGGAGCTTCTCTCTCTCTCT 939  
Db 914 TCGCCGTCGTTGTTTCCGGGAGAAGTTTCCGGAGAGAAGGTGCTCTCTCTCTCTCTCT 973  
Qy 940 CCCTTTGGGGCTTTGCTCTTCTTACTTTTATGGTGAGATAAAGTCTGGCGAGGA 991  
Db 974 CTCCTTTGGGATTTGCTCTTACTTCTACGCGGAGTTTAAATCCGGCAAGAA 1025  
RESULT 8  
AF078531 1227 bp mRNA linear PLN 20-APR-2000  
LOCUS Arabidopsis thaliana purine permease (PUP1) mRNA, complete cds.  
DEFINITION Arabidopsis thaliana (thale cress)  
ACCESSION AF078531 GI:7620006  
VERSION AF078531.1  
KEYWORDS  
SOURCE  
ORGANISM Arabidopsis thaliana  
Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.  
REFERENCE 1 (bases 1 to 1227)  
AUTHORS Gillissen,B., Burkle,L., Andre,B., Kuhn,C., Rentsch,D., Brandl,B.  
and Frommer,W.B.  
TITLE A new family of high-affinity transporters for adenine, cytosine,  
and purine derivatives in Arabidopsis  
JOURNAL Plant Cell 12 (2), 291-300 (2000)  
MEDLINE 20129770  
PUBMED 10662864  
REFERENCE 2 (bases 1 to 1227)  
AUTHORS Gillissen,B., Andre,B., Rentsch,D., Buerkle,L., Kuehn,C., Brandl,B.  
and Frommer,W.B.  
TITLE Direct Submission  
JOURNAL Submitted (15-JUL-1998) Institut fuer Botanik, University of  
Tuebingen, Auf der Morgenstelle 1, Tuebingen, Baden-Wuerttemberg  
72076, Germany  
FEATURES  
source  
1..1227  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/cultivar="Columbia"  
/db\_xref="taxon:3702"  
gene  
1..1227  
/gene="PUP1"  
CDS  
27..1097  
/gene="PUP1"  
/note="AtPup1"  
/codon\_start=1  
/product="purine permease"  
/protein\_id="AA564547.1"  
/db\_xref="GI:7620007"  
/translation="MNGLLIINCIIITGTCGGPLLRLLRLLYFNGGKRINWPMFLSTA  
GFPIILVLSFLSRNRGNPNNAENRKLFLMETPLFIASIVIGLLTGLDNL  
YSYGLAYLPVSTSLIIGTQAFNALFAILLKQKTFPSINAVLLTVGIGIALHS  
DGAPKESKKEYVGVFLMTVAALLYAFLLPLVELTYKAROBITPLVLEIOMVNC  
LAATFCVIGMFIVGPKVTAREAREBFKIGSVFYVALIVITGIIWGFELGATGIVF  
CASLASGVLISVLLPVTVEFAVVCVREKQAEKVSLLLSLWGFVSFYFGEPKSGKK  
VVDKPPPEPPELPLPVSDIYA"  
ORIGIN  
Query Match 40.1%; Score 428.4; DB 8; Length 1227;  
Best Local Similarity 66.4%; Pred. No. 4.6e-102;  
Matches 672; Conservative 0; Mismatches 316; Indels 24; Gaps 3;  
Qy 1 AGACAAGATGGTGAGGCTCTTGTGATCATATAAATCTGCATAATCTTACCCATAGAAACT 60  
Db 19 AGAAGAATGAAGAATGGTTTGATAATCATATAAATCTGTAATCTCTACTATAGAACAT 78  
Qy 61 GTGGAGGTCTTTGATTATGCGTCTCTACTTCAACCAATGGCGGTAAGAGGATTTGGTTCT 120  
Db 79 GTGGAGGTCTTTGTTAACTGCTCTACTTCAACCAATGGCGGAAACGAATCTGGTTCA 138  
Qy 121 CTACGTTTCTTGAACACTGCAGGCTTTCTCTGTATCTTCTCTCTCTCTCTCTCTCTTACA 180





mRNA	genefinder"	
CDS	<5251..>.5466	
	/gene="At2g333690"	
	5251..5466	
	/gene="At2g333690"	
	/codon_start=1	
	/product="hypothetical protein"	
	/protein_id="AAC69143.1"	
	/db_xref="GI:2253578"	
repeat_region	/translation="MSKSEKQELPLETSPYTKYEDIEYKKNAGTSGHQDVKPGHG	
gene	GOTDAPTPSGDAFSAIDSANQAKK"	
	complement(6459..6492)	
	/rpt_family="GA)n"	
	7252..9045	
	/gene="At2g33700"	
	/note="synonym: Ti1B8.2"	
mRNA	join(<7252..7434,7522..7713,7864..8069,8275..9045)	
CDS	/gene="At2g33700"	
	join(7252..7434,7522..7713,7864..8069,8275..8836)	
	/gene="At2g33700"	
	/codon_start=1	
	/product="putative protein phosphatase 2C"	
	/protein_id="AAC69126.1"	
	/db_xref="GI:1707015"	
	/translation="MSMDFSPLLTVLEGDFNKNDSATETDLENLDDTROISKGP	
	PRHLTSSATRLQANADVCLNLMKSLDDKSELPVYRSGSCAEQAKQFMEDEHI	
	CIDDLVHLGAAIQSSIGAPYGVDFGCGTDAHFVRKNILRFIVEDSSPFLCKKA	
	IKSAFLKADYFADDSLDISGTTALTAFIFGRRLIIANAGDCRAVLRRGRAIEIS	
	KDHKPNCAEVRIEKLGQVYDVGVLQSLVARAIGDMHMKPKGSACPLSPPELQ	
	EDLSEDDFEITMGCDLGMVSSQCAVTIARKELMHNDPERGCSRELVRALKRNTC	
	DNLTVIVCFSPDPQRIEIRMSRVRSSISAEGLLKLLKGLVDGVP"	
repeat_region	complement(11621..11649)	
gene	/rpt_family="AT rich"	
	complement(12323..13704)	
	/gene="At2g33710"	
	/note="synonym: Ti1B8.3"	
mRNA	complement(join(<12323..12798,13524..>13704))	
CDS	/gene="At2g33710"	
	complement(join(12323..12798,13524..13704))	
	/gene="At2g33710"	
	/codon_start=1	
	/product="putative AP2 domain transcription factor"	
	/protein_id="AAC69127.1"	
	/db_xref="GI:1707016"	
	/translation="MHSCKRPLSPESMAGNREKELCCSTLSSESVDVFSBLTGQ	
	FIPSIDQSSSLTQESNKRQYRQRPWKAAEIRDPNKAARVWLGTFTDA	
	BEAALAYDKAAPFPGHKAKLNPPEHIRVNPOLYPSPATSHDRIIVTPPSPPPIAP	
	DILLDQYHFQSRSDSSANLSMMLSSSSSLNHQGLRPNLEDGENVKNISIHRRK	
repeat_region	complement(14124..14156)	
repeat_region	/rpt_family="AT rich"	
gene	complement(14389..14425)	
	/rpt_family="TA)n"	
	16828..18041	
	/gene="At2g33720"	
	/note="synonym: Ti1B8.29; predicted by genscan and	
	genefinder"	
	join(<16828..17139,17373..>18041)	
mRNA	/gene="At2g33720"	
CDS	join(16828..17139,17373..18041)	
	/gene="At2g33720"	
	/codon_start=1	
	/product="hypothetical protein"	
	/protein_id="AAC69145.1"	
	/db_xref="GI:2253580"	
	/translation="MKMPFPFASAKTQYLPHDESSENKSKSLVSTTSLSCENPNKR	
	KXNSDEVINI SCIDRYKLQTOVERKILARMNLISYQKAEDEWYGVSTELTFKDPWII	
	KKVPHFASVLDMAPNSVNTCHLLDTPESPENABESLSLDCIFDYDKTWHPDNPVANK	
	PTSBEAFLNANWMTSKARKEEKYVSTELTLTVADPHTLAKWMTKSSIGNLYRL	
	VLSKAFVDIHLIRYLPDDQMWVEDSGLAVEVDHDTSDVHNLALKKWAQSFVILV	
	SGWRKCFVDRGLQGVIGMWDORSKSLHFCVLSRSETNWDSEAPLPSP"	
gene	complement(18731..20932)	
	/gene="At2g33730"	

mRNA	/note="synonym: Ti1B8.4"	
CDS	complement(<18731..>20932)	
	/gene="At2g33730"	
	complement(18731..20932)	
	/gene="At2g33730"	
	/codon_start=1	
	/product="putative U5 small nuclear ribonucleoprotein, an	
	RNA helicase"	
	/protein_id="AAC69128.1"	
	/db_xref="GI:1707017"	
	/translation="MKRTINEVVGSTSVNTLDSIAKKPVLTKKQREELAKRRQDQ	
	ISEORVREQLRPEPETEDVSGDNTNRDKDRDRDRDRDRDRDRDRDRDRDRDR	
	DRDRDRERDRERDRDRERDRERDRERDRERDRERDRERDRERDRERDRER	
	DAIKQYLGGKKPKKRVIRPSEKFRFSFDWENTEDTSRDNVLYQNHEAQLLGRGF	
	RAGMRREQKQAQKHEKMRDEIRKKGIVKEPPEAAQVRBEAADTYDSFDMRVD	
	RHWSDKLEEMTERDWIRIFREDFNISYKGSRIPRMRSWEESKLTSLMLKAVERAGYK	
	KPSDIPMAAIPLGLQQRDVIGIAFTGSKTAAFLVPLMAYISRLPMPSENETEGPYA	
	VMAPTRELAAQIBETVKFAHLYGFRVTSIVGGOSIEOGLKLTQCEIVIAATPGR	
	IDCLERRYAVLNQCNVYVLEADRMIDMGPEQVAGVLDAMPSSNLKPENEEELDEK	
	KIYRTTMSATMPPGYERLARKYLRNPVVTIGTAGTITDILISQHVIMKSESKFR	
	LQKLDELGETAIVFVNTKNCDSIAKNLDKAGYRVITLHGKSKQSQREISLSEGFRA	
	KRYNVLATDVVGRGIDIPDAHVINYDMPKHIEMYTHRIGTRAGKSGVATSFLLT	
	HDTEVFYDLKQMLVQSNASVPELARHEASRFKEGTVPDRPPRHSDTVYN"	
repeat_region	complement(20588..20718)	
gene	/rpt_family="C(CAGAGA)n"	
	complement(21321..22493)	
	/gene="At2g33735"	
	/note="synonym: Ti1B8.17; supported by full length cDNA:	
	Ceres: 269036"	
mRNA	complement(join(<21321..21655,21744..21814,22051..22126,	
	22221..22298,22384..>22493))	
	/gene="At2g33735"	
CDS	complement(join(21584..21655,21744..21814,22051..22126,	
	22221..22298,22384..22446))	
	/gene="At2g33735"	
	/codon_start=1	
	/product="Expressed protein"	
	/protein_id="AAM15518.1"	
	/db_xref="GI:20198310"	
	/translation="MWCSDMGDSHGDFDFEDYKDYKHYKLELNCDSADDEIRSSFIR	
	LALKWHPDKFEEDSATSRFQEIENAYQVLSPIARQYDKRMRIYEHNTTELLNEY	
	KELILTGNLGLGVKHYLW"	
repeat_region	22524..22544	
	Query Match 30.9%; Score 330; DB 8; Length 114144;	
	Best Local Similarity 67.0%; Pred. No. 4.9e-76;	
	Matches 468; Conservative 0; Mismatches 230; Indels 0; Gaps 0;	
Qy	1 AGACAAGAATGGTCAAGGCTCTTGTGATCATAAATGCATAATTTCTAGCCATAGGAACT 60	
Db		
	26319 AGATGAAGATGAAGACAGTTCCTTGTAATCATAACTGTATATCTTGGCCATTTGGAAC 26260	
Qy	61 GTGAGGTGCTTTTGATTATCGCTCTTACTTCAACAATGGCGGTAAAAGGATTTGGTTCT 120	
Db		
	26259 GTGAGGCGCTCTAATGATGCTCTTACTTCCAAATGGTGGGAAAGGATCTGGTTTC 26200	
Qy	121 CTAGCTTTCTTGAACGAGGCTTTCCCTGTATCTTCTCATCTCCTCTCTCTCTTACA 180	
Db		
	26199 CAAGCTTCTTCAACCGTTGGTTGTCCACTATTTTCTTCCCTCTTCTTATCTTCTCC 26140	
Qy	181 TTACCCGGCGCAGAACAAATGTGGTGATAGTACAAAGTTTCTTCTTATCAACCGC 240	
Db		
	26139 TCCGCCCTGCTGTTGGCTTGAAGAACAGGATCCATTTTCTCATGAAACCTC 26080	
Qy	241 GTCTTCTTATCGCGCTGTTATTGTAGGCATTTCTCAGGGTTTGATAAATCTTGTATG 300	
Db		
	26079 CTCTCTTATCGCGCTATCGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 26020	
Qy	301 CATATGTATAGCTTATCTCCAGTTTCTACAGTGTCTTATCATCTTACAGTTTACAGTAG 360	
Db		
	26019 CTTACGGGTTAGCTTATATCCCTGTTTCTACTCGCTCTTTGATCATCTCCGCGCAATTAG 25960	
Qy	361 CTTTATAGCTATCTTCTCATTTCTTATGTTTAACATAAAAGTTTCACTCTCTTTTACCATCA 420	





```
Db      435  GTGCTGGGCTCTCACCAGCGCGACGACTTCGCTACGCTACGGCCCTCGCCCTACCTG 494
Qy      321  CAGATTTCACAGCTGCTTATCTATGCTTACAGTTCAGTTCAGTTCAGTTCAGTTCAGTTC 380
Db      495  CGGCTGTCACCTCGCCATCTCTATCTCCAGCAGCTGCGCTTCCACCGTCTTCTTCGCG 554
Qy      381  TTCTTCATGTTAAACATAAGTTCACTCTCTTTTACCATCAATGCTGTGTGTGTGTGACT 440
Db      555  TGCCTCATGTCGGCAGCGCTCACCGGCGGCGCTGAACGCGCTGGCGCTGCTCACT 614
Qy      441  GTTGCTGTCGGGTTTGGGAATGCATACCGAACTGATAAGCCAGTTCATGAGCACTCAC 500
Db      615  ATCGCGCGCGTGTGCTGGCGCTCCACGCGTCCAAAGGACCGACCGCGCGGGTGACCA 674
Qy      501  AAGCAGTACATAACTGTTCTTCTGATTACTGTAGCAGCAGCTGTATGTATGCTTTGATC 560
Db      675  GGAAGTACTGATGGCTTCTTCTCTCACCTCGCGCGCGCGCTGTAGCGGCTCATC 734
Qy      561  TTGCCATTAGTGGAACTTGCTTTACCAGAAAGCTAAG-----CAAACCATGAGCTAT 611
Db      735  CTGCGCTGCTGAGCTCGGTACAGCAGCGCGCGCGCGCGCGCGCTGACGTAC 794
Qy      612  ACCCTGTGCTCAGTTCAGTTGATTTGTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 671
Db      795  GCGCTGCTGATGGAGATGAGCTGCTCATGGCTTCTTCTCGCCACCGCTTCTGTACCG 854
Qy      672  GGTATGTTATCATCGCTGTTGATTTCAAGCAGCGCTTACCAGAGCGCTTACCAGAGCAG 731
Db      855  GGTGTGATGCTGTAACAAGACTTC---CAGGCGATCCCGAGGAGCGGCAAGCAGTAC 911
Qy      732  CTTGGAGAGGCAATTTCTTATGTGCTGCTGTGTTTTCAGGCATCATATGCGAAGGCTTC 791
Db      912  CTGGGTGAAGCGCGTACTAGCTGTGTGTTTCAACGCGCTCTCTGGGAGTTCTTTC 971
Qy      792  TTCCTGGAGGCAATTTGATTTCTTCCACATCGTCTCTGCTCTCGGCTCGGCTTATGATA 851
Db      972  TTCGTGGGCGCGTGGCGTCTATCTTCTGCTGACACGCTGCTCGCGGATCATCATC 1031
Qy      852  TCAGTGTCTTTTGCCAAATACAGAGTTTACGTGTATATCTACCATGAAGTTTCAA 911
Db      1032  GCGGTGTTATCCCGATCAGGAGTGCTGGCGTCACTTCTTCTCCAGAGAGTTACG 1091
Qy      912  GCTGAGAAGGCACTTCTCTTCTCTCTCCCTTTGGGCTTGTCTCTTCTTCTTATGGT 971
Db      1092  AGCGAGAAGGCGTGGCGCTGCTCTGCTCTGCGGCGCTCTGCGCTCTCTCTACCGC 1151
Qy      972  GAGATAAAGTCTGGCGAGGATAAAGGAGA 1001
Db      1152  GAGTACGCGCGAGCCCAAGCCCAAGAGAAA 1181
```

```
RESULT 11
AK066610
LOCUS   AK066610
DEFINITION
Oryza sativa (japonica cultivar-group) cDNA clone:J013073M15, full
insert sequence.
ACCESSION
AK066610
VERSION
FLI_CDNA; CAP trapper.
KEYWORDS
Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoidae; Oryzaceae; Oryza.
REFERENCE
1
AUTHORS
The Rice Full-Length cDNA Consortium, National Institute of
Agrobiological Sciences Rice Full-length cDNA Project Team,
Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K.,
Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I.,
Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C.,
Ohtsuki,K., Shishiki,T., Foundation of Advancement of International
Science Genome Sequencing & Analysis Group; Ohtsuki,K., Murakami,K.,
```

```

Iida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsunoda,Y.,
Kurosaki,T., Kodama,T., Masuda,H., Kobayashi,M., Xie,Q., Lu,M.,
Narikawa,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Niikura,J.,
Ikeda,R., Ishibiki,J., Kawamata,M., Yoshimura,A., Miura,J.,
Kusumegi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., RIKEN:
Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Arakawa,T., Fukuda,S.,
Hara,A., Hashidume,W., Hayatsu,N., Imotani,K., Iehi,Y., Itoh,M.,
Kagawa,I., Kondo,S., Konno,H., Miyazaki,A., Osato,N., Oka,Y.,
Sato,R., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,
Yoshino,M. and Hayashizaki,Y.
Collection, mapping, and annotation of over 28,000 cDNA clones from
japonica rice
Science 301 (5631), 376-379 (2003)
22752273
12869764
2 (bases 1 to 1505)
Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Doi,K.,
Fujimura,T., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W.,
Hayashida,K., Hayashizaki,Y., Hayatsu,N., Hiramoto,K., Hiraoka,T.,
Hori.F., Hotta,I., Iida,J., Iida,Y., Ikeda,R., Inamura,K.,
Imotani,K., Ishibiki,J., Ishii,Y., Ishikawa,M., Itoh,M., Kogawa,I.,
Kanagawa,S., Katoh,H., Kawagashira,N., Kawai,J., Kawamata,M.,
Kikuchi,S., Kishikawa-Hirozane,T., Kishimoto,N., Kobayashi,M.,
Kodama,T., Kojima,K., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Kurosaki,T., Kusumegi,T., Li,C., Lu,M.,
Masuda,H., Matsubara,K., Matsuyama,T., Miura,J., Miyazaki,A.,
Mizuno,K., Murakami,K., Murata,M., Nagata,T., Nakamura,M.,
Namiki,T., Narikawa,R., Niikura,J., Nishi,K., Nomura,K.,
Numasaki,R., Ohneda,E., Ohno,M., Ohtsuki,K., Oka,M., Ooka,H.,
Osato,N., Oka,Y., Ohtsuki,K., Ohtsuki,K., Ohtsuki,K., Sakai,K.,
Sakazume,N., Sano,H., Sasaki,D., Sato,K., Satoh,K., Shibata,K.,
Shinagawa,A., Shiraki,T., Shishiki,T., Sogabe,Y., Sugano,S.,
Sugiyama,A., Suzuki,K., Suzuki,Y., Tagami,M., Tagami-Takeda,Y.,
Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A.,
Toyata,T., Tsunoda,Y., Ueda,M., Waki,K., Xie,Q., Yahagi,W.,
Yamada,H., Yamamoto,M., Yasunishi,A., Yazaki,J., Yokomizo,S. and
Yoshimura,A.
Direct Submission
Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of
Agrobiological Sciences, Department of Molecular Genetics, Head of
Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki
305-8602, Japan (E-mail:ekikuchi@nias.affrc.go.jp)
Tel:81-29-838-7007, Fax:81-29-838-7007)
This clone is one of the 28K full-length cDNA clones from japonica
rice.
URL : http://cdna01.dna.affrc.go.jp/cDNA/
NIAS Rice Full-length cDNA Project Team: Kikuchi,S., Satoh,K.,
Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J.,
Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T.,
Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T. and
Yamamoto,M.
FAIS Genome Sequencing & Analysis Group: Ohtsuki,K., Iida,Y.,
Fujimura,T., Ikeda,R., Ishibiki,J., Kawamata,M., Kobayashi,M.,
Kodama,T., Kurosaki,T., Kusumegi,T., Lu,M., Masuda,H., Miura,J.,
Mizuno,K., Narikawa,R., Niikura,J., Oka,M., Ryu,R., Sugano,S.,
Sugiyama,A., Suzuki,Y., Tsunoda,Y., Ueda,M., Xie,Q., Yokomizo,S.,
Yoshimura,A., Matsubara,K. and Murakami,K.
Genome Exploration Research Group in Riken Genomic Sciences Center
and Genome Science Laboratory in Riken: Adachi,J., Aizawa,K.,
Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hanagaki,T.,
Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K.,
Hiraoka,T., Hori,F., Iida,J., Inamura,K., Imotani,K., Ishii,Y.,
Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawai,J.,
Kishikawa-Hirozane,T., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Ohno,M., Osato,N.,
Oca,Y., Satoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H.,
Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F.,
Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toyota,T., Waki,K.,
Yasunishi,A. and Hayashizaki,Y.
Location/Qualifiers
1. .1505
FEATURES
source
```

```
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="J013073M15"
```

## ORIGIN

Query Match	21.6%	Score 230;	DB 8;	Length 1505;
Best Local Similarity	54.4%;	Pred. No. 1.1e-49;		
Matches 539;	Conservative 0;	Mismatches 430;	Indels 21;	Gaps 3;
Qy	21	CTTTGTGATCATAAACTGCATAATTTCTAGCCATAGAAACTGTGAGGTCTCTTTGATTATG	80	
Db	198	CTTTGGCGCTCAACTGTGGGATCTGTGTGTGGCCACCAGGCGGGCGCTCATCAGC	257	
Qy	81	CGTCTCTACTTCAACAATGGCGGTAAAGAGATTTGGTTCTCTACGTTTCTTGAACATGCA	140	
Db	258	CGCTCTACTTTCAGCAAGGGCGGCATAGGAGTGGCTCTCGGGTGGCTCCAGACCGCC	317	
Qy	141	GGCTTTCTGTATCTTCATTCCTCTGCTCTCTCTTACATTAACCCGGCGGAGAAGCAAC	200	
Db	318	GGGTGGCGCTGTCTGTGTGCCGTGGCGGCTCTACCTCAGCCGCGCGCGAGAC	377	
Qy	201	AATGTGGGTGATAGTACAAGTTTCTTTCTTATCAAAACGGGTCTTCTTATGCGCGTGT	261	
Db	378	-----GCGCGCGCGCGCTGTTTCTCAAGCGGACGGCGTGTCTGGCGGGGTC	428	
Qy	261	ATTGTAGGCATTTCTCTCAGGGTTTGATACTTGTATGATCATGATAGTATATCTT	320	
Db	429	GGGCTCGGGTTCTCAATGGCGCGAGACTTTATATACGCTTACGGCTCGCCTACTCG	488	
Qy	321	CCAGTTTCTACAGCTGCTTATCAATTCCTTTCACAGTTTAGTCTTTATAGTATCTTCTCA	380	
Db	489	CCGGTGTCCACTCCGCGCATCTCATCTCCAGCGAGCTGGGTTTCCCGTCTTCTTCGCG	548	
Qy	381	TTCTTCATGGTTAAATAGTTTCACTCTCTTTTACCATCAATGCTGTGTGTGTTGACT	440	
Db	549	TGCTCTCATCTGTGGCGACGGCTTTCACCGCGCGACGCTGAACGCGGTGGCGCTCACT	608	
Qy	441	GTTGGTCTGCGGTTTTGGGAATGCATAACGAAACTGATAGCCAGTTTCATGACACTCAC	500	
Db	609	ATCGCGCGCTGTGTCTGGGCTGCAACGGTCCAAGGACCGCCCGCGGGTTACCAGT	668	
Qy	501	AAGCAGTACATACTGGTTTCTTGATTACTGTGTAGCAGAGCTGTATGTATGCTTTCATC	560	
Db	669	GGGAAGTACTGGATGGGCTTCTTCTTCCATCCCTCGGCGCGCGGCTGTACGGGCTCATC	728	
Qy	561	TTGCGCATTTAGTGGAACTTGCTTACAGAAAGCTAAG-----CAAAACCATCAGCTAT	611	
Db	729	CTGCGCTGCTGAGCTCGGTGTAACAACGCGCGGGAGCGCGCGCTCAGGTAC	788	
Qy	612	ACCTTGTGCTCGAGTTTCCAGTTGATTTTGTGTCCTTGTCTTATGTCAGCGTCAATC	671	
Db	789	CGCTGTGTGATGGAGATGCAGCTGGTCATGGGCTTCTTGGCCACCGGCTTCTGACCGCTC	848	
Qy	672	GGTATGTTCACTCGTGGTGTATTTCAACGAGCCTTACCAAGAGCAAGAGAGTTCAAG	731	
Db	849	GGCATGTGCTGTAACAAGGACTTC--CAGCGGATCCCGAGGGAGGCGAAGCAGTACGAG	905	
Qy	732	CTTGGAGAGGCAATGTTCTATGTGGTGGCTGTGTTTTTTCAGCCATCATATGCAAGGCTTC	791	
Db	906	CTGGGTGAAGCGGGTACTACGTGTGTGTGTTTCAAGCGGCTCTGTGGAGTCTCTTC	965	
Qy	792	TTCTTTGGAGGCCATGGATTAATCTTCTCCACATCGTCTCTCGTCTCGGGTATTAATGATA	851	
Db	966	TTCTGTGGCGCGCTCGCGCTCATCTTCTGGGTGACACGCTGTCTCGCGGGATCATCATC	1025	
Qy	852	TCAGTGTCTTTGCCAAATTACAGAGGTTTTAGCTGTATTATCTACCATGAAGATTTCAA	911	
Db	1026	GCCGTGTTTCATCCGATCAAGGAGTGTGGGCGTCACTCTCTCCACAGAGAGTTTCAGC	1085	
Qy	912	GCTGAGAAAGGAGCTTTCTCTTGCTCTCTCCCTTTTGGGCTTTGTCTCTTACTTTTATGGT	971	

Db	1086	AGCGAAGGGCGTGGCGCTCTCGCTCTGGGGCCCTCGCCTCCTACTCTACGGC	1145
Qy	972	GAGATAAAGTCTGGCGAGGATAAAAGGAGA	1001
Db	1146	GAGTACGCCGACGCCAAGGCCCAAGAGAAA	1175
RESULT 12			
AK0641135			
LOCUS			
DEFINITION	Oryza sativa (japonica cultivar-group)	cdna clone:002-102-G04, full	
	AK0641135	1363 bp	mRNA
ACCESSION			
VERSION	AK064135.1	GI:32974153	
KEYWORDS	FLI cDNA; oligo-capping.		
SOURCE	Oryza sativa (japonica cultivar-group)		
ORGANISM	Oryza sativa (japonica cultivar-group)		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.		

RESULT 12  
AK064135

LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

REFERENCE  
AUTHORS

**TITLE**

JOURNAL  
MEDLINE

## REFERENCE AUTHORS

JOURNAL

COMMENT

The Rice Full-Length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-Length cDNA Project Team,  
Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K.,  
Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H.,  
Ohtsuka, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C.,  
Kojima, K., Shishiki, T., Foundation of Advancement of International  
Science Genome Sequencing & Analysis Group: Ootomo, Y., Murakami, K.,  
Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y.,  
Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M.,  
Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Nikura, J.,  
Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J.,  
Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN:  
Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S.,  
Hara, A., Hashikume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M.,  
Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Osato, N., Ota, Y.,  
Saio, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,  
Yoshino, M. and Havaaphizaki, Y.

ROSANO, M. and NAYANIZAKI, I.  
 Collection, mapping, and annotation of over 28,000 cDNA clones from  
 japonica rice  
 Science 301 (5631), 376-379 (2003)  
 22752273  
 12869764

2 (bases 1 to 1363)

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K., Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hizamoto, K., Hiraoka, T., Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imamura, K., Ikehata, I., Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M., Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Kurosaeki, T., Kusumegi, T., Li, C., Lu, M., Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A., Mizuno, K., Murakami, K., Murata, M., Negata, T., Nakamura, M., Namiki, T., Narikawa, R., Nishikura, J., Nishi, K., Nomura, K., Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H., Osato, N., Ota, Y., Otomo, Y., Ryu, R., Saitho, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K., Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and Yoshimura, A.

Direct Submission  
Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of  
Agrobiological Sciences, Department of Molecular Genetics, Head of  
Laboratory of Gene Expression: 2-1-2 Kannondai, Tsukuba, Ibaraki  
305-8602, Japan (E-mail: skkuchi@nias.affrc.go.jp,  
Tel:81-29-838-7007, Fax:81-29-838-7007)  
This clone is one of the 28K full-length cDNA clones from japonica  
rice.



Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Hotta, I., Iida, Y., Ikeda, R., Imamura, K., Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M., Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koyasu, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M., Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A., Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M., Namiki, T., Nariikawa, R., Niikura, J., Nishi, K., Nomura, K., Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H., Ootani, R., Ota, Y., Otonari, Y., Ryo, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K., Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and Yoshimura, A.

# TITLE JOURNAL

Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail: skikuchi@nias.affrc.go.jp, Tel: 81-29-838-7007, Fax: 81-29-838-7007)

This clone is one of the 28K full-length cDNA clones from japonica rice.

URL: <http://cdna01.dna.affrc.go.jp/cDNA/>  
NIAS Rice Full-Length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T. and Yamamoto, M.

FAIS Genome Sequencing & Analysis Group: Otonari, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Nariikawa, R., Niikura, J., Oka, M., Ryo, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K. and Murakami, K.  
Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Osato, N., Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K., Yasunishi, A. and Hayashizaki, Y.

## FEATURES

source  
1. .1476  
Location/Qualifiers  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="mRNA"  
/cultivar="Nipponbare"  
/db\_xref="taxon:39947"  
/clone="J023087G23"

## ORIGIN

Query Match 17.1%; Score 182.2; DB 8; Length 1476;  
Best Local Similarity 50.9%; Pred. No. 4.7e-37;  
Matches 458; Conservative 0; Mismatches 438; Indels 3; Gaps 1;  
1  
63 GGAGGCTCTTGGTATGCTCTCTACTTCAACAATGGCGGTAAGGATTTGGTCTCT 122  
123 ACCTTTCTTGAACGACGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 182

Db 364 ACATTCGTCACAACTGCTGGCTTTCCGATTTTGTTCATTTGCCCTATTCTTTTCCATTCA 423  
OY 183 ACCGCGCGCAGAGCAACAATGTGGGTGATAGTACAAAGTTTCTTTCTTATCAAAACGCGT 242  
Db 424 AAGACATCTTCTACACAAACAGTCATAGTAGTCTTGGCCCTTACAAATTTCTATTCGCCAAA 483  
OY 243 CTTCCTTATCGCGCTGTTATTGTAGGCATTTCTCTCAGGGTTTGATAACTACTTTGTATGCA 302  
Db 484 ATTACTCTGATATATGTTCTCTGGCCCTCATCATTTGCTGCAGACGACTTGGTGTATTC 543  
OY 303 TATGGTATAGCTATCTTCCAGTTTCTACAGCTGCTCTTATCATTTGCTTTCACAGTTAGCT 362  
Db 544 TATGGCTACTATATCTTCCGGTCTCAACATATTCGCTCATCTGTGCTAGTCAGCTTGGC 603  
OY 363 TTTATAGCTATCTTCTCATTTCTTCTGTTTAAACATAAAGTTCACCTCTTTTACCATCAAT 422  
Db 604 TTTCAATGCTCTTCTCATATTTCTCAATGCTCAAAATTTACCCCTCTGATTTTCAAT 563  
OY 423 GCTGTGTGTTGTTGACTGTTGCTGCGGTTTGGGAATGCTATACCGAAACTGATAAG 482  
Db 664 TCGGTAGTCTCTTACGTTTCTGCTTCACTCTTGGAGTTGATGAAGATTCTCAGGGA 723  
OY 483 CAGTTTCATGACACTCACAAGCAGTACATAACTGTTTCTTCTTCTTCTTCTTCTTCTTCT 542  
Db 724 ACTACTAGTATATATCATGCGGAAGTACATTTTGGGTTTCTCTGTGACACTAGGGGATCA 783  
OY 543 GTTATGTATGCTTTCATCTTGCATTTAGTGAACCTTGTTCACAGAAAGCTAAGCAACC 602  
Db 784 GCTACATCTCCTCATTTCTCTCTGATGCAAGTGCACATTTGAGAAGGTTATTAAGAGG 843  
OY 603 ATGAGCTATPACCTTTGTGCTCGAGTTTCCAGTTTGTGTTGCTCTCTCTCTCTCTTCTT 662  
Db 844 GAGACCTTCTCAGTTTGTGTTGAACATGACAGATATATACAGCTCTCTGTGCAACATTTGG 903  
OY 663 ACGGTATCGGTATGTTTCATCCTCGTGTGATTTCAAGAGCGCTTACCAAAAGAGCAAGA 722  
Db 904 TCTCTTGTGGGTATTTTCAAGTGTGTAAGTGAATGA ---CTTTCAAGGAGAGATGTCAT 960  
OY 723 GAGTTCAAGCTTGGAGAGCATTTGTTTATGTGTGGTGTGTTTTCAGCCCATCATATGS 782  
Db 961 GCATTTCCAGTCTGGAAGCTGTCATATGTAATGACACTGCTGTGAGCGGCTATATCTTGG 1020  
OY 783 CAAGGCTTCTTCTTGGGAGCCATTGGATTAATTTCTTCACATCGTCTCTCGTCTCGGCT 842  
Db 1021 CAGGTAGCATCATGTTGGAGTGTGGGATTTGATCTTTGTGTGTCATCGTCTTCTTCAAT 1080  
OY 843 ATTATGATATCAGTCTTTTGGCAATATACAGAGGTTTGTAGCTGTTATATTTACCATGAA 902  
Db 1081 GTGATAAGCACCTAGCTCTACCCCATCATTTCTCTGTTTGTGCTGTGATTTTCTTCTATG 1140  
OY 903 AAGTTTCAAGCTGAGAGGAGGACTTTCTCTTGTCTCTCTCTCTCTCTCTCTCTCTCTCT 961  
Db 1141 AAGATGGATGGAGTAAGATTAAGTATAGCTATGCTGATGGCAATTTGGGATTTATGTCATA 1199

## RESULT 14

AK072660  
LOCUS  
DEFINITION  
Oryza sativa (japonica cultivar-group) cDNA clone:J023134011, full insert sequence.  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
The Rice Full-Length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-Length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Hotta, I.,

Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group, Ohtsuki, K., Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y., Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M., Nariakawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J., Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J., Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN, Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Osato, N., Oka, Y., Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Yoshino, M., and Hayashizaki, Y.

Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice

Science 301 (5631), 376-379 (2003)

12752273

12869764

2 (bases 1 to 1617)

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K., Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imamura, K., Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M., Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M., Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A., Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M., Namiki, T., Nariakawa, R., Niikura, J., Nishi, K., Nomura, K., Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H., Osato, N., Ota, Y., Ohtsuki, K., Ohtsuki, K., Ohtsuki, K., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K., Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S., and Yoshimura, A.

Direct Submission

Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of Agricultural Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression, 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail: skikuchi@nias.affrc.go.jp, Tel: 81-29-838-7007, Fax: 81-29-838-7007)

This clone is one of the 28K full-length cDNA clones from japonica rice.

URL : <http://cdna01.dna.affrc.go.jp/cDNA/>

NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T., and Yamamoto, M.

FAIS Genome Sequencing & Analysis Group: Ohtsuki, S., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Nariakawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K., and Murakami, K.

Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Itoh, M., Kagawa, I., Iida, J., Imamura, K., Imotani, K., Ishii, Y., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Osato, N., Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K.,

Yasunishi, A. and Hayashizaki, Y.  
Location/Qualifiers  
1..1617  
/organism="Oryza sativa (japonica cultivar-group)"  
/molecule="mRNA"  
/cultivar="Nipponbare"  
/db\_xref="taxon:39947"  
/clone="J023134011"

#### ORIGIN

Query Match 17.1%; Score 182.2; DB 8; Length 1617;  
Best Local Similarity 50.9%; Pred. No. 4.7e-37;  
Matches 458; Conservative 0; Mismatches 438; Indels 3; Gaps 1;

QY 63 GGAGGTCTTTGATATGCGTCTCTACTTCAACAATGGCGGTAAAGGATTTGGTTCTCT 122  
DB 459 GCATCTACACTTCTTGGGAGGTTCTACTACAATCAAGGCGGCAATAGCAAGTGGATGTCC 518  
QY 123 ACCTTTCTTGAACCTGCGAGGCTTTCTCTGTTATCTTCAATTTCTCTGCTCTCTCTTACATT 182  
DB 519 ACATTCGTCCAACTGCTGGCTTTCCGATTTTGTTCATTGGCCCTATTTCTTTTCCATTCA 578  
QY 183 ACCCGGCGGAGAACCAACAATGTGGGTGATAGTACAAGTTTCTTTCTTATCAAAACCGCGT 242  
DB 579 AAGACATCTCTACACAAACAGTCACTAGTAGTCTCTGCCCTACAAATTTCTATCCCAA 638  
QY 243 CTTCTTATCGCGCTGTATGTAGGCAATCTCTCAGGTTTGTATACTACTTCTGTATGCA 302  
DB 639 ATTACTCTGATATATGTGTCTGGGCTCATCATTTCTGCAGACGACTTGTATGTATTC 698  
QY 303 TATCGTATAGCTTATCTTCCAGTTTCTACAGTCTCTCTTATCATTTCTTACAGTTAGCT 362  
DB 599 TATGCGCTACTATATCTTCGGTCTCAACATATTCGGTCACTCTGTGTAGTCACTTGGC 758  
QY 363 TTTATAGCTATCTTCTCATTTCTCATGGTTTAAACATAAAGTTTCACTCTTTTACCATCAAT 422  
DB 759 TTCAATGCTGCTCTCTCATATTTCTCAATGCTCAAAAATTTACCCCTCTGATTTTCAAT 818  
QY 423 GCTGTTGTTGTGTAGTCTGTTGGTCTGCGGTTTGGGAATGTCATACCGAACTGATAAG 482  
DB 819 TCCGTAGTCTCTCTTACGTTTCTTGTCTCACTCTCTGGAGTTGATGAAGATTCTTCAGGGA 878  
QY 483 CCAGTTTATGAGACTCAACAGCAGTACATAAAGTGTGTTTCTTGTATCTGTAGCAGCAGCT 542  
DB 879 ACTACTAGTATATACATGGAGTACATTTTGGTTTCTGTTGACACTAGGGGCATCA 938  
QY 543 GTTATGATGCTTTTCATCTTGGCCATAGTGAACCTCTCTTACCAAGAAAGCTTAAGCAAC 602  
DB 939 GCTACATCTCGCTCATCTCTCTCCCTGATCAAGTCAATTTGAGAAGGTTATTAAGAGG 998  
QY 603 ATGAGCTATACCTTGGCTCGAGTTCAGTTGATTTTGTGTCCTCTCTCTTCTTATTTGTC 662  
DB 999 GAGACCTTCTCAGTTGTTGAAACATGACATATATACAGCTCTCTGTTGGCAACATTTGGCT 1058  
QY 663 AGCGTCTATCGTATGTTTCATCGCTGTTGATTTCAAGCAGGCTTTACCAAAAAGCAAGA 722  
DB 1059 TCTCTTGTGGTTATTTGCAAGTGGTGAATGATGA--CTTTACAGAGGAGATGCAT 1115  
QY 723 GAGTTCAAGCTTGAGAGGCAATTTCTATGTGTGTGGCTGTGTTTTCAGCAATCATATGG 782  
DB 1116 GCATTTCCAGTCTGGAAGCTGTATATGTAATGACACTGCTGTGGAGGGCTATATCTTGG 1175  
QY 783 CAGGCTTCTTCTTGGGAGCCATTGGATTAATCTTCTCCACATCGTCTCTCGTCTCGGT 842  
DB 1176 CAGTAGCATCAGTTGGAGTGGTGGGATGATCTTTGTGGTGCATCGCTGTTTTCAAAT 1235  
QY 843 ATTATGATATCAGTCTCTTTTGGCCAAATPACAGAGGTTTATCTGTTATATCTACCATCAA 902  
DB 1236 GTGATAGACCCCTAGCTCTACCCATCATTTCTCTGTTTGTGCTGTGATTTCTTTTCATGAC 1295  
QY 903 AAGTTTCAAGCTGAGAGGAGCAATTTCTCTTGTGCTCTCTCCCTTTGGGGCTTTGTCTCTTA 961  
DB 1296 AAGATGATGGAGTAAAGATTATAGCTATGCTATGCTGATGGCCATTTGGGGAATTTATGTCATA 1354









GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 1, 2004, 02:40:28 ; Search time 626.985 Seconds  
(without alignments)  
9996.736 Million cell updates/sec

Title: US-09-913-767-5  
Perfect score: 1194  
Sequence: 1 tcatgagatataataaacat.....tgtgatcaagcatatttcc 1194

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues 8269772  
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N Geneseq\_23Sep04:\*  
1: geneseqn1980s:\*  
2: geneseqn1990s:\*  
3: geneseqn2000s:\*  
4: geneseqn2001as:\*  
5: geneseqn2001bs:\*  
6: geneseqn2002as:\*  
7: geneseqn2002bs:\*  
8: geneseqn2003as:\*  
9: geneseqn2003bs:\*  
10: geneseqn2003cs:\*  
11: geneseqn2003ds:\*  
12: geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1190.8	99.7	1194	3	Aaa97923 A. thalia
2	1015.8	85.1	3387	6	ABZ14407 Arabidops
3	1015.8	85.1	3387	8	ADA68246 Arabidops
4	813.4	68.1	1293	3	Aaa97922 A. thalia
5	459.8	38.5	592	10	ABX56692 Arabidops
6	405.8	34.0	1071	3	Aaa97925 A. thalia
7	369.6	31.0	1047	6	ADG88190 A. thalia
8	369.6	31.0	1081	3	Aaa97924 A. thalia
9	317.2	26.6	1145	3	Aaa97921 A. thalia
10	316.8	26.5	1292	3	AAC51119 Arabidops
11	316.8	26.5	1295	3	AAC37289 Arabidops
12	305.8	25.6	492	3	AAC36845 Arabidops
13	270	22.6	2660	3	AAC47974 Arabidops
14	215.4	18.0	83698	6	ABN85767 Arabidops
15	213.2	17.9	2175	8	ADA70748 Rice gene
16	212.8	17.8	1152	8	ADA70213 Rice gene
17	196	16.4	1403	3	AAC37652 Arabidops
18	135.4	11.3	1049	3	Aaa97920 A. thalia
19	134.4	11.3	1155	8	ADA69686 Rice gene
20	120	10.1	386	6	ABQ85215 Arabidops
21	118.4	9.9	332	12	ADP92185 Cotton ex

22	111.8	9.4	1225	3	Aaa97919
23	111	9.3	2586	3	AAC44184 Arabidops
24	103.2	8.6	1040	8	ADA69568 Rice gene
c 25	75	6.3	5520	5	AAC68924 FLR1 rece
26	73.2	6.1	1173	12	ADN72520 Thale cre
27	66.2	5.5	446	3	AAC37188 Arabidops
28	65.8	5.5	1421	3	AAC39779 Arabidops
29	64.6	5.4	477	3	AAC37273 Arabidops
30	64.2	5.4	1418	3	AAC46395 Arabidops
31	62.6	5.2	774	6	ABN99127 Arabidops
32	51.6	4.3	1200	8	ADA70627 Rice gene
c 33	50.6	4.2	2000	8	ADA71938 Rice gene
34	43.6	3.7	493	3	AAC36831 Arabidops
35	42.2	3.5	2000	8	ADA71938 Rice gene
36	40.4	3.4	34901	10	ADC86940 Human GPC
c 37	37.4	3.1	358	6	ABV98933 Human pan
38	36.8	3.1	369	4	Aak80520 Human imm
39	36.8	3.1	369	4	Aak80518 Human imm
40	36.8	3.1	369	4	Aak80519 Human imm
41	36.8	3.1	370	4	Aak61600 Human imm
c 42	36.8	3.1	6741	3	AAA10595 Gene enco
c 43	36.4	3.0	651	10	ADF57711 Human pol
c 44	36.4	3.0	2928	3	AAC39967 Arabidops
c 45	36.4	3.0	4590	5	Aah24065 Yeast AOD

ALIGNMENTS

RESULT 1

AAA97923  
ID AAA97923 standard; DNA; 1194 BP.

AC AAA97923;

DT 19-JAN-2001 (first entry)

DE A. thaliana PUP1 DNA #5.

KW PUP1; transgenic plant; nucleobase transporter; apical dominance;  
KW flowering behaviour; senescence; pesticide distribution; ds.

OS Arabidopsis thaliana.

PN DE19907209-A1.

PD 24-AUG-2000.

PF 19-FEB-1999; 99DE-01007209.

PR 19-FEB-1999; 99DE-01007209.

XX (FROM/) FROMMER W.

XX Gillissen B, Buerkle L, Andre B, Frommer WB;

XX WPI; 2000-566202/53.

PT Nucleic acid, useful for producing transgenic plants with altered  
PT nucleobase transport, encodes a nucleobase transporter protein of  
PT Arabidopsis thaliana.

PS Claim 1f; Page 14; 24pp; German.

CC This invention describes a novel nucleic acid encoding a plant nucleobase  
CC transporter (I). (I) is produced by complementation of a nucleobase  
CC transport (NBT)-defective host cell with a plant gene bank by selection  
CC of NBT-positive cells. (I) is used to isolate homologous sequences from  
CC bacteria, fungi, plants, animals and humans, for expression of the  
CC encoded protein (II) in prokaryotic or eukaryotic cells, for inhibiting  
CC expression of (II) (when in antisense orientation), and to produce  
CC transgenic crop plants. The transgenic plants have modified nucleobase  
CC transport properties, e.g. altered affinity and substrate specificity

CC that may result in more efficient nucleobase transport in leaves, changes  
CC in apical dominance, flowering behaviour and senescence, or improved  
CC distribution of pesticides. This sequence encodes the Arabidopsis  
CC thaliana PUP1 protein which is described in the method of the invention  
XX  
SQ Sequence 1194 BP; 298 A; 279 C; 231 G; 386 T; 0 U; 0 Other;

Query Match	99.7%;	Score 1190.8;	DB 3;	Length 1194;
Best Local Similarity	99.8%;	Pred. No. 0;		
Matches 1192; Conservative	0;	Mismatches 2;	Indels 0;	Gaps 0;
Qy 1	TCATGAGATATAATAAAACATGAGTCTTAATTTTTCAGGTGACACAGAACTTTAGAAAGCAAC	60		
Db 1	TCATGAGATATAATAAAACATGAGTCTTAATTTTTCAGGTGACACAGAACTTTAGAAAGCAAC	60		
Qy 61	CTTTATAGATCATGAGTGGTAACTGAAATCATCATCAGCTGTCCTCAAAACCGAGAAC	120		
Db 61	CTTTATAGATCATGAGTGGTAACTGAAATCATCATCAGCTGTCCTCAAAACCGAGAAC	120		
Qy 121	TATAAAGGTGGCTTCGTGTCCTCATATACGTAATCTTTGTCCTCTTTTGGCAGCCACTA	180		
Db 121	TATAAAGGTGGCTTCGTGTCCTCATATACGTAATCTTTGTCCTCTTTTGGCAGCCACTA	180		
Qy 181	GCTACAAATTTCTGGGTAGATTGACTATGAAATGGAGAAATAGCACATATGTGGTAAACA	240		
Db 181	GCTACAAATTTCTGGGTAGATTGACTATGAAATGGAGAAATAGCACATATGTGGTAAACA	240		
Qy 241	CTTCTTCAACTCATTTGGGCTTCCCTGTACTGGTTCTGTTCCGCTTCTTTTCTCGAATCAGG	300		
Db 241	CTTCTTCAACTCATTTGGGCTTCCCTGTACTGGTTCTGTTCCGCTTCTTTTCTCGAATCAGG	300		
Qy 301	CAACCCAAATCAACAGATACAAATTTCACTGAGTCCCTTCCCTTACACACCCTTGCATCG	360		
Db 301	CAACCCAAATCAACAGATACAAATTTCACTGAGTCCCTTCCCTTACACACCCTTGCATCG	360		
Qy 361	GTTTACTGTGCACTGGAGCTGCTAGTGTCCGCTTATGCTTATTTGTCTGCAGTAGGGTTG	420		
Db 361	GTTTACTGTGCACTGGAGCTGCTAGTGTCCGCTTATGCTTATTTGTCTGCAGTAGGGTTG	420		
Qy 421	CTCTACTACAGTCTCTACTTTTCCCTCATCTTGGCCTCACAGTTGGCTTCACTGCC	480		
Db 421	CTCTACTACAGTCTCTACTTTTCCCTCATCTTGGCCTCACAGTTGGCTTCACTGCC	480		
Qy 481	TTTTTCTCATATTTCTTAACTCGCAAAGTTCACTCCTTTGATAGTCACTTTCTTTCGTT	540		
Db 481	TTTTTCTCATATTTCTTAACTCGCAAAGTTCACTCCTTTGATAGTCACTTTCTTTCGTT	540		
Qy 541	CTCCTCAGTGATCTCTGCTCTTCTTGTGCTCAACACTGATTCAGAAAATCAACTAAT	600		
Db 541	CTCCTCAGTGATCTCTGCTCTTCTTGTGCTCAACACTGATTCAGAAAATCAACTAAT	600		
Qy 601	GTATCTAGAGTACAGTATGTGATCGGGTTCAATATGTACCATCGGTGCTCCGCTGGGATT	660		
Db 601	GTATCTAGAGTACAGTATGTGATCGGGTTCAATATGTACCATCGGTGCTCCGCTGGGATT	660		
Qy 661	GGACTGTATCTATCTCTGATACAAATGCTCTTTCAGGAAAGTTTTCAGAAAGCATATCC	720		
Db 661	GGACTGTATCTATCTCTGATACAAATGCTCTTTCAGGAAAGTTTTCAGAAAGCATATCC	720		
Qy 721	TCAGCAGTACCGACTTGGCCATTTCACAGTCTCTAGTTGCGAGTTGTGTAGTTCTCAT	780		
Db 721	TCAGCAGTACCGACTTGGCCATTTCACAGTCTCTAGTTGCGAGTTGTGTAGTTCTCAT	780		
Qy 781	GGACTTTTTGCAAGTGGAGTGGGAAACTTTGCCAAGTGAGATGAGAAACTACAAACTC	840		
Db 781	GGACTTTTTGCAAGTGGAGTGGGAAACTTTGCCAAGTGAGATGAGAAACTACAAACTC	840		
Qy 841	GGGAAAGTCTCATATGTTTTCAGCTTTTACGCTCGGCAGCTATTCCTGGCAGGTACACT	900		
Db 841	GGGAAAGTCTCATATGTTTTCAGCTTTTACGCTCGGCAGCTATTCCTGGCAGGTACACT	900		
Qy 901	CCTGGTCTTGTGGGATGATCTTCAGTCACTCTGTTGTTCTCAATTCCTAATCAGCT	960		

Db	901	CTTGGTCTTGTTGGGATTGATCTTCGAGTCACTCTCTGTGTCTTCCAAATCCATAACAGCT	960
Qy	961	GTGGGATTGCTATAGTTCAGTTCCAGTTTCGGCAGTGATAGTTCATGATAGATGGACGCA	1020
Db	961	GTGGGATTGCTATAGTTCAGTTTCGGCAGTGATAGTTCATGATAGATGGACGCA	1020
Qy	1021	TCCAAATCTTCTCCATATATTTAGCTATCTGGGGCTCTTTCATTTCTGCTATCAGCAC	1080
Db	1021	TCAAAATCTTCTCCATATATTTAGCTATCTGGGGCTCTTTCATTTCTGCTATCAGCAC	1080
Qy	1081	TACCTCCAGCAAGAGATTGAATACCTAGCCACACAAGTCTGTAGGAGATCTTCATCTA	1140
Db	1081	TACCTCCAGCAAGAGATTGAATACCTAGCCACACAAGTCTGTAGGAGATCTTCATCTA	1140
Qy	1141	CCTGTTGAGGAAGTTCACACAAACATACAAAGTGTGTGATCAAAAGCATATTTCC	1194
Db	1141	CCTGTTGAGGAAGTTCACACAAACATACAAAGTGTGTGATCAAAAGCATATTTCC	1194
RESULT 2			
ID	ABZ14407	standard; DNA; 3387 BP.	
XX	AC	ABZ14407;	
XX	DT	21-JAN-2003 (first entry)	
XX	DE	Arabidopsis thaliana stress regulated gene SEQ ID NO 2212.	
XX	KW	Arabidopsis thaliana; plant; gene; stress; transgenic; ds.	
XX	OS	Arabidopsis thaliana.	
XX	PN	WO200216655-A2.	
XX	PD	28-FEB-2002.	
XX	PF	24-AUG-2001; 2001WO-US026685.	
XX	PR	24-AUG-2000; 2000US-0227866P.	
XX	PR	26-JAN-2001; 2001US-0264647P.	
XX	PR	22-JUN-2001; 2001US-0300111P.	
XX	PA	(SCRI ) SCRIPPS RES INST.	
XX	PA	(SYGN ) SYNGENTA PARTICIPATIONS AG.	
XX	PI	Harper JF, Kreps J, Wang X, Zhu T;	
XX	DR	WPI; 2002-304127/34.	
XX	PT	Identifying a stress condition to which a plant cell has been exposed and	
XX	PT	producing plants with increased tolerance to these abiotic stresses.	
XX	PS	Claim 144; SEQ ID NO 2212; 577pp + Sequence Listing; English.	
XX	CC	The invention relates to identifying a stress condition to which a plant	
XX	CC	cell has been exposed, comprising: (a) contacting nucleic acid	
XX	CC	representative of expressed polynucleotides in the plant cell with an	
XX	CC	array or probes representative of the plant cell genome; and (b)	
XX	CC	detecting a profile of expressed polynucleotides in the plant cell	
XX	CC	characteristic of a stress response. The method is useful in the	
XX	CC	production of transgenic plants, cells and seeds and in producing plants	
XX	CC	with increased tolerance to abiotic stress. The present sequence is that	
XX	CC	of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used	
XX	CC	in methods of the invention. Note: The sequence data for this patent is	
XX	CC	not represented in the printed specification but is based on sequence	
XX	CC	information supplied to Derwent by the European Patent Office	
XX	SQ	Sequence 3387 BP; 870 A; 781 C; 654 G; 1082 T; 0 U; 0 Other;	
Query Match 85.1%; Score 1015.8; DB 6; Length 3387;			
Best Local Similarity 95.1%; Pred. No. 1.1e-300;			
Matches 1081; Conservative 0; Mismatches 2; Indels 54; Gaps 1;			

QY 38 GTGACGAGAACTTAGAAGCAAAACCTTATAGATCATGAGGTGGTAACTGAATCATCATCAT 97  
Db 1205 GTGACGAGAACTTAGAAGCAAAACCTTATAGATCATGAGGTGGTAACTGAATCATCATCAT 1264  
QY 98 CAGCTGTGCTCAAAACCGAGAACTATAAAGGTGGCTTCGTGCTCCATATAGTAACTCT 157  
Db 1265 CAGCTGTGCTCAAAACCGAGAACTATAAAGGTGGCTTCGTGCTCCATATAGTAACTCT 1324  
QY 158 TTGTCTCTTTTGGCCAGCCACTAGCTACAAATCTGGGTAGATTGTAATGAAATGGAG 217  
Db 1325 TTGTCTCTTTTGGCCAGCCACTAGCTACAAATCTGGGTAGATTGTAATGAAATGGAG 1384  
QY 218 GAAATAGCATAATGTGGTAACACTTCTTCAACTCAATTTGGCTTCCCTGTACTGTTCTGT 277  
Db 1385 GAAATAGCATAATGTGGTAACACTTCTTCAACTCAATTTGGCTTCCCTGTACTGTTCTGT 1444  
QY 278 TCCGCTTCTTTTCTCGAATCAGCAACCCAAATCAACAGATACAAATTTTCAGTCAGTCCC 337  
Db 1445 TCCGCTTCTTTTCTCGAATCAGCAACCCAAATCAACAGATACAAATTTTCAGTCAGTCCC 1504  
QY 338 CTTTCTTCCACCCCTTGATCGGTTTACTTTGTGCACTGGACTGTAGTGTCCGCTTATG 397  
Db 1505 CTTTCTTCCACCCCTTGATCGGTTTACTTTGTGCACTGGACTGTAGTGTCCGCTTATG 1564  
QY 398 CTTATTGTCTGCAGTAGGGTTGCTCTACTTACCAGTCTCTACTTTTCTCCCTCATCTGG 457  
Db 1565 CTTATTGTCTGCA----- 1578  
QY 458 CCTCAGATGGCCCTTCACTGCGCTTTTCTCATATATTTCTTAATCTGCAAAAGTTCACTC 517  
Db 1579 -----TTGGCCCTTCACTGCGCTTTTCTCATATATTTCTTAATCTGCAAAAGTTCACTC 1630  
QY 518 CTTTGTATGTCAGTCTTTGCTCTCTCACTGTATCTCTGCTCTTCTTGTGGTCAACA 577  
Db 1631 CTTTGTATGTCAGTCTTTGCTCTCTCACTGTATCTCTGCTCTTCTTGTGGTCAACA 1690  
QY 578 CTGATTTCAGAAACTCAACTAATGTATCTAGAGTACAGATGTGATCGGGTTTCATATGA 637  
Db 1691 CTGATTTCAGAAACTCAACTAATGTATCTAGAGTACAGATGTGATCGGGTTTCATATGA 1750  
QY 638 CCATCGGTGCTCGCTGGATGGAGTGTACTATCTCTGATACAAATGCTCTTCAGGA 697  
Db 1751 CCATCGGTGCTCGCTGGATGGAGTGTACTATCTCTGATACAAATGCTCTTCAGGA 1810  
QY 698 AAGTTTTTCAGAAAGCATACCTCTCAGCAGTCAGGACTTGGCCATTTACCAGTCTCTAG 757  
Db 1811 AAGTTTTTCAGAAAGCATACCTCTCAGCAGTCAGGACTTGGCCATTTACCAGTCTCTAG 1870  
QY 758 TTGCGAGTTGTGTAGTTCTCATAGGACTTTTTCGAAAGTGGAGTGGGAACTTTTGCCAA 817  
Db 1871 TTGCGAGTTGTGTAGTTCTCATAGGACTTTTTCGAAAGTGGAGTGGGAACTTTTGCCAA 1930  
QY 818 GTGAGATGAGAAACTCAAACTCGGAAAGTGTATATGTTTGTAGCTTTAGCTCGGCAG 877  
Db 1931 GTGAGATGAGAAACTCAAACTCGGAAAGTGTATATGTTTGTAGCTTTAGCTCGGCAG 1990  
QY 878 CTAATTTCTGGCAAGTCTACACTCTGCTGTTGGGATTGATCTTCGAGTCACTCTCTG 937  
Db 1991 CTAATTTCTGGCAAGTCTACACTCTGCTGTTGGGATTGATCTTCGAGTCACTCTCTG 2050  
QY 938 TGTCTTCCAAATTCAGCTGTGGGATTGGCTATAGTTCCAGTTGGCGGAGTGATAG 997  
Db 2051 TGTCTTCCAAATTCAGCTGTGGGATTGGCTATAGTTCCAGTTGGCGGAGTGATAG 2110  
QY 998 TTTTTCATGATGAATGGAGCATCCAAATCTTCTCCATTAATTTTGTAGCTATCTCGGCT 1057  
Db 2111 TTTTTCATGATGAATGGAGCATCCAAATCTTCTCCATTAATTTTGTAGCTATCTCGGCT 2170  
QY 1058 TCCCTTTTCATTCGCTATCAGCACTACCTCGAGAAAGAGTTGAATACTAGCCACAA 1117  
Db 2171 TCCCTTTTCATTCGCTATCAGCACTACCTCGAGAAAGAGTTGAATACTAGCCACAA 2230

QY 1118 GTGCTGTAGGAGATCTTCACTTACCTGTGAGGAAGGTCAACAACATACAAAGTG 1174  
Db 2231 GTGCTGTAGGAGATCTTCACTTACCTGTGAGGAAGGTCAACAACATACAAAGTG 2287  
RESULT 3  
ID ADA68246 standard; DNA; 3387 BP.  
XX ADA68246;  
XX 20-NOV-2003 (first entry)  
XX Arabidopsis thaliana gene, SEQ ID 908.  
XX Plant; bacterial infection; fungal infection; viral infection; ds.  
XX Arabidopsis thaliana.  
XX WO2003000898-A1.  
XX 03-JAN-2003.  
XX 22-JUN-2001; 2001WO-IB001105.  
XX 22-JUN-2001; 2001WO-IB001105.  
XX (SYGN ) SYNGENTA PARTICIPATIONS AG.  
XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;  
XX Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;  
XX WPI; 2003-175290/17.  
XX Identifying at least one gene involved in plant resistance or response to  
XX pathogenic infection for conferring resistance or tolerance to a plant to  
XX bacterial, fungal or viral infection by determining or detecting plant  
XX gene expression.  
XX Claim 6; SEQ ID NO 908; 899pp; English.  
XX The present invention relates to a method (M1) for identifying genes  
XX involved in plant resistance or response to pathogenic infection. M1  
XX comprises identifying a gene whose expression is significantly altered in  
XX the incompatible interaction of plant gene expression relative to  
XX expression of the gene in an uninfected plant, in a mutant plant that  
XX does not express a gene associated with response to pathogenic infection,  
XX or in a corresponding incompatible or compatible interaction. (M1) is  
XX useful for conferring resistance to resistance or tolerance to a plant to  
XX bacterial, fungal or viral infection. The present sequence was used to  
XX illustrate the invention.  
SQ Sequence 3387 BP; 870 A; 781 C; 654 G; 1082 T; 0 U; 0 Other;  
Query Match 85.1%; Score 1015.8; DB 8; Length 3387;  
Best Local Similarity 95.1%; Pred. No. 1.1e-300;  
Matches 1081; Conservative 0; Mismatches 2; Indels 54; Gaps 1;  
QY 38 GTGACGAGAACTTAGAAGCAAAACCTTATAGATCATGAGGTGGTAACTGAATCATCATCAT 97  
Db 1205 GTGACGAGAACTTAGAAGCAAAACCTTATAGATCATGAGGTGGTAACTGAATCATCATCAT 1264  
QY 98 CAGCTGTGCTCAAAACCGAGAACTATAAAGGTGGCTTCGTGCTCCATATAGTAACTCT 157  
Db 1265 CAGCTGTGCTCAAAACCGAGAACTATAAAGGTGGCTTCGTGCTCCATATAGTAACTCT 1324  
QY 158 TTGTCTCTTTTGGCCAGCCACTAGCTACAAATCTGGGTAGATTGTAATGAAATGGAG 217  
Db 1325 TTGTCTCTTTTGGCCAGCCACTAGCTACAAATCTGGGTAGATTGTAATGAAATGGAG 1384  
QY 218 GAAATAGCATAATGTGGTAACACTTCTTCAACTCAATTTGGCTTCCCTGTACTGTTCTGT 277  
Db 1385 GAAATAGCATAATGTGGTAACACTTCTTCAACTCAATTTGGCTTCCCTGTACTGTTCTGT 1444

QY 278 TCCGCTTCTTTCTCGAATCAGGCACCCAAATCAACAGATACAAATTTTCAGTCAGTCCC 337  
Db 1445 TCCGCTTCTTTCTCGAATCAGGCACCCAAATCAACAGATACAAATTTTCAGTCAGTCCC 1504  
QY 338 CTTCTCTTCCACACCCCTTCGATCGGTTTACTTGTGCACTGGAGCTGCTAGTGTCCGCTTATG 397  
Db 1505 CTTCTCTTCCACACCCCTTCGATCGGTTTACTTGTGCACTGGAGCTGCTAGTGTCCGCTTATG 1564  
QY 398 CTTATTTCTGCTGCACTAGGGTGTCTACTTACCAGTCTCTACTTTCTCCCTCATCTTGG 457  
Db 1565 CTTATTTCTGCA----- 1578  
QY 458 CCTCACAGTTGGCCTTCACTGCTTTTCTCATATTTCTTAACTCGCAAAAGTTCATC 517  
Db 1579 -----TTGGCCTTCACTGCTTTTCTCATATTTCTTAACTCGCAAAAGTTCATC 1630  
QY 518 CTTTGATAGTCAGTCTTTTGTCTTCCTCACTGCTATCTCTGCTCTTCTTGTGCTCAACA 577  
Db 1631 CTTTGATAGTCAGTCTTTTGTCTTCCTCACTGCTATCTCTGCTCTTCTTGTGCTCAACA 1690  
QY 578 CTGATTCAGAAAACTCAACTAATATCTAGTACAGTACAGTATGTGATCGGGTTTCATATGTA 637  
Db 1691 CTGATTCAGAAAACTCAACTAATATCTAGTACAGTACAGTATGTGATCGGGTTTCATATGTA 1750  
QY 638 CCATCGGTGCTTCGCTGGATTGGAGCTTACTATCTCTGATACAAATGCTCTCAGGA 697  
Db 1751 CCATCGGTGCTTCGCTGGATTGGAGCTTACTATCTCTGATACAAATGCTCTCAGGA 1810  
QY 698 AAGTTTTCCAGAAATACATCCTCAGCAGTCAAGGCTTGGCCATTTACCAGTCTCTAG 757  
Db 1811 AAGTTTTCCAGAAATACATCCTCAGCAGTCAAGGCTTGGCCATTTACCAGTCTCTAG 1870  
QY 758 TTGCGAGTTGTAGTCTCTAGGACTTTTGTGCAAGTGGAGAGTGGAAACTTTTGCAA 817  
Db 1871 TTGCGAGTTGTAGTCTCTAGGACTTTTGTGCAAGTGGAGAGTGGAAACTTTTGCAA 1930  
QY 818 GTGAGATGAGAACTTACAACTCGGAAAGTGTCTATGTTTTCATCTTACCTTCGGCAG 877  
Db 1931 GTGAGATGAGAACTTACAACTCGGAAAGTGTCTATGTTTTCATCTTACCTTCGGCAG 1990  
QY 878 CTTATTTCTGGCAAGTCTACACTCTGCTGCTTGTGGGATTGATCTTCAGTCACTCCTG 937  
Db 1991 CTTATTTCTGGCAAGTCTACACTCTGCTGCTTGTGGGATTGATCTTCAGTCACTCCTG 2050  
QY 938 TGTCTTCCAAATCCATAACAGCTGTGGGATTGCCCTATAGTTCAGTTCGCGCAGTGATAG 997  
Db 2051 TGTCTTCCAAATCCATAACAGCTGTGGGATTGCCCTATAGTTCAGTTCGCGCAGTGATAG 2110  
QY 998 TTTTCCATGATAGATGAGCGCATCCAAATCTTCTCCATTTATTTTAGCTATCTCGCGCT 1057  
Db 2111 TTTTCCATGATAGATGAGCGCATCCAAATCTTCTCCATTTATTTTAGCTATCTCGCGCT 2170  
QY 1058 TCCTTTCATTCGCTCTATCAGCACTACCTCGACGAAAGAGTTGAAATAGTACAGCACACAA 1117  
Db 2171 TCCTTTCATTCGCTCTATCAGCACTACCTCGACGAAAGAGTTGAAATAGTACAGCACACAA 2230  
QY 1118 GTGCTGTAGGAGATCTTCACTACCTGTGTGAGGAGGTCAACAAACATACAAAGTG 1174  
Db 2231 GTGCTGTAGGAGATCTTCACTACCTGTGTGAGGAGGTCAACAAACATACAAAGTG 2287

RESULT 4

AAA97922 standard; DNA; 1293 BP.

ID AAA97922

XX

AC AAA97922;

XX

DT 19-JAN-2001 (first entry)

XX

DE A. thaliana PUP1 DNA #4.

XX

KW PUP1; transgenic plant; nucleobase transporter; apical dominance;

flowering behaviour; senescence; pesticide distribution; ds.

Arabidopsis thaliana.

DE19907209-A1.

24-AUG-2000.

19-FEB-1999; 99DB-01007209.

19-FEB-1999; 99DE-01007209.

(FROM/) FROMMER W.

Gillissen B, Buerkle L, Andre B, Frommer WB;

WPI; 2000-566202/53.

Nucleic acid, useful for producing transgenic plants with altered

nucleobase transport, encodes a nucleobase transporter protein of

Arabidopsis thaliana.

Claim 1f; Page 13; 24pp; German.

This invention describes a novel nucleic acid encoding a plant nucleobase transporter (I). (I) is produced by complementation of a nucleobase transport (NBT)-defective host cell with a plant gene bank by selection of NBT-positive cells. (I) is used to isolate homologous sequences from bacteria, fungi, plants, animals and humans, for expression of the encoded protein (II) in prokaryotic or eukaryotic cells, for inhibiting expression of (II) (when in antisense orientation), and to produce transgenic crop plants. The transgenic plants have modified nucleobase transport properties, e.g. altered affinity and substrate specificity that may result in more efficient nucleobase transport in leaves, changes in apical dominance, flowering behaviour and senescence, or improved distribution of pesticides. This sequence encodes the Arabidopsis thaliana PUP1 protein which is described in the method of the invention

Sequence 1293 BP; 345 A; 288 C; 247 G; 413 T; 0 U; 0 Other;

Query Match 68.1%; Score 813.4; DB 3; Length 1293;  
Best Local Similarity 81.6%; Pred. No. 9.6e-239;  
Matches 969; Conservative 0; Mismatches 206; Indels 12; Gaps 2;  
Qy 20 TGAGTGTAAATTTTTCAGGTGACCGAGAACTAGACGAAACCACTATAGATCATGAGTGG 79  
Db 107 TGTGTTTGTATGTTTTCAGATCATAAACAAATAGAGCAAAACCTTAACAGGTGAGGAGAA 166  
Qy 80 TAACT-----GAATCATCATCATCAGCTGCTCAACCGAGAACTATAAAAGGT 130  
Db 167 TGAATACCATGGAATCGAATCTCGTCCGTACCTCAATCGAAGAACTATAAGAAAT 226  
Qy 131 GGCTTCGTGCTCCATATACGTAATCTTTGCTCTTTTGGCAGCCACTAGTCAATATC 190  
Db 227 GGCTTCGTATTCATTTACGTGTTCTTGTCTTGTGCTGCTGCAAGCACTTTCTACAATTT 286  
Qy 191 TGGTAGATTTGACTATGAAAATGGAGAAATAGCAATATGTTGTAACACTTCTTCAAC 250  
Db 287 TGGGAGAGTTTACTATGAAAATGGTGGGAAAGATACATGATGGGAAACACTTGTCCAAC 346  
Qy 251 TCATTGGCTTCCCTGTACTGTTCTGTTCCGCTTCTTTTCTCGAATCAGGCAACCCAAAT 310  
Db 347 TAATCGGCTTCCCTGTGTTCTGTTCTTCTGCTTCTTTTCCCAACCAAAAATCCCAAC 406  
Qy 311 CAACAGATACAAATTTTCAGTTCAGTCCCTTCTTCCACCACTTGCATATCGGTTTACTTGT 370  
Db 407 CAACAGAGCAGATTTTCAGAAAGTTCTCTTCTTCCATCTTGTGATCAGTTTACATCG 466  
Qy 371 GCATGAGCTAGTGTGCTCGCTTATGCTTATTTGCTGAGTAGGTTGCTCTACTTAC 430  
Db 467 TTAGTGAGATATTAGTGTCTGCTAACTCTTATATGCTCTGTTGGTTTACTACTTAC 526  
Qy 431 CAGTCTCTACTTTTCTCCCTCATCTTGGCCTCACAGTTGGCCTTCACTGCTTTTCTCAT 490



Qy 217 GGAAATAGACATATGTGGTAACACTTCTTCAACTCATTTGGCTTCCCTGTACTGTTCTG 276  
|||  
Db 226 GGGAAAGACATATGTGGTAACACTTCTTCAACTCATTTGGCTTCCCTGTACTGTTCTG 285  
|||  
Qy 277 TTCCGCTTCTTTTCTCGAATCAGGCAACCCAAATCAACAGATACAAAATTTTCAGTCAGTCC 336  
|||  
Db 286 TTCCGCTTCTTTTCTCGAATCAGGCAACCCAAATCAACAGATACAAAATTTTCAGTCAGTCC 345  
|||  
Qy 337 CCTTCCTTACACACCTTGCATCGGTTTACTTGTGCACTGGACTGCTAGTGTCCGCTTAT 396  
|||  
Db 346 CCTTCCTTACACACCTTGCATCGGTTTACTTGTGCACTGGACTGCTAGTGTCTGCTTAT 405  
|||  
Qy 397 GCTTATTTGTCTGCAGTAGGTTGCTCTACTTACCAGTCTCTACTTTCCTCCTCATCTTG 456  
|||  
Db 406 GCTTATTTGTCTGCAGTTGGGTTGCTTTACTTACCAGTCTCTACTTTCCTCCTCATCTTG 465  
|||  
Qy 457 GCCTCACAGTTGGCCTTCACTGCCTTTTCTCATATTTTCTTAACTCGCAAAAGTTCAC 516  
|||  
Db 466 GCCTCACAGTTGGCCTTCACTGCCTTTTCTCATATTTTCTTAACTCGCAAAAGTTCAC 525  
|||  
Qy 517 CCTTGTATGATGAGTCTTGTGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 576  
|||  
Db 526 CCTTGTATGATGAGTCTTGTGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 585  
|||  
Qy 577 ACTGATT 583  
|||  
Db 586 ACTGATT 592  
|||

## RESULT 6

AAA97925

ID AAA97925 standard; DNA; 1071 BP.

XX

AC AAA97925;

XX

19-JAN-2001 (first entry)

XX

A. thaliana PUP1 DNA #7.

XX

PUP1; transgenic plant; nucleobase transporter; apical dominance;

XX flowering behaviour; senescence; pesticide distribution; ds.

XX

OS Arabidopsis thaliana.

XX

PW DE19907209-A1.

XX

24-AUG-2000.

PD

19-FEB-1999; 99DE-01007209.

XX

19-FEB-1999; 99DE-01007209.

XX

(FROM/) FROMMER W.

XX

Gillissen B, Buerkle L, Andre B, Frommer WB;

XX

WPI; 2000-566202/53.

XX

Nucleic acid, useful for producing transgenic plants with altered

XX nucleobase transport, encodes a nucleobase transporter protein of

PT Arabidopsis thaliana.

XX

Claim 2; Page 15; 24pp; German.

PS

This invention describes a novel nucleic acid encoding a plant nucleobase

XX transporter (I). (I) is produced by complementation of a nucleobase

CC transport (NBT)-defective host cell with a plant gene bank by selection

CC of NBT-positive cells. (I) is used to isolate homologous sequences from

CC bacteria, fungi, plants, animals and humans, for expression of the

CC encoded protein (ii) in prokaryotic or eukaryotic cells, for inhibiting

CC expression of (ii) (when in antisense orientation), and to produce

CC transgenic crop plants. The transgenic plants have modified nucleobase

XX transport properties, e.g. altered affinity and substrate specificity

CC that may result in more efficient nucleobase transport in leaves, changes  
CC in apical dominance, flowering behaviour and senescence, or improved  
CC distribution of pesticides. This sequence encodes the Arabidopsis  
CC thaliana PUP1 protein which is described in the method of the invention  
XX  
SQ Sequence 1071 BP; 307 A; 236 C; 208 G; 320 T; 0 U; 0 Other;

Query Match 34.0%; Score 405.8; DB 3; Length 1071;  
Best Local Similarity 64.5%; Pred. No. 1.4e-113;  
Matches 624; Conservative 0; Mismatches 337; Indels 6; Gaps 1;

Qy 164 TCTTTTGGCCAGCCACTAGCTACAAATCTGGGTAGATTGTACTATGAAAATGGAGAAATA 223  
|||  
Db 47 TCATTGGCCAAATCAGTTGCTACAAATCTGGGAGACTATACTATGAAAATGGAGAAACA 106  
|||  
Qy 224 GCATATATGTTGTAACACTTCTTCAACTCAATGGCTTCCCTGACTGTTGTTCTGTTCCGCT 283  
|||  
Db 107 GCAATGGCTAGCAACGGTAGTTTTCAGCTTGTAGGCTTTCCTATTTCTACTTCCATATCATC 166  
|||  
Qy 284 TCTTTTCTCGAATCAGGCAACCCAAATCAACAGATACAAATTTTCAGTCAGTCCCTTCCT 343  
|||  
Db 167 TCTTGTGAGTCA-----AAACACATACAACTCAGAGAGATGGCAAATTAACCTCAC 220  
|||  
Qy 344 TCACCACCCCTTGCATCGGTTTACTTGTGCACCTGGACTGTAGTGTCCGCTTATGCTTATT 403  
|||  
Db 221 TTAGGAACCGTGCAATTAGTTTACATAGTGTCTGGACTTCTTGTAGGAGCAGCTTGCTACC 280  
|||  
Qy 404 TGCTGCAGTAGGTTGCTCTACTTACCAGTCTCTACTTTTCTCCCTCATCTTGGCCTCAC 463  
|||  
Db 281 TATATTCCATTTGGACTGCTTTTACCTACCTGTTTCTACCCCTTCCCTGATGTGTCATCAC 340  
|||  
Qy 464 AGTTGGCTTCACTGCTTTTCTCATATTTTCTTAACCTCGCAAAAGTTCACCTCTTTGA 523  
|||  
Db 341 AGTTAGCCCTTACCCGCTTCTTCTTATTTACTCAACTCACAATAACTTACTTCTCTATCA 400  
|||  
Qy 524 TAGTCAGTCTTCTTGTCTTCTCCTCACTGTATCTCTGTCTTCTTGTGGTCAACACTGATT 583  
|||  
Db 401 TTTTGAATCTCTTTTCTCTCTCTCATATATCTTCCACCCCTCTTGCAATTAATAACGAGG 460  
|||  
Qy 584 CAGAAAACCTCAACTAATGTATCTAGAGTACAGTATGTATCGGGTTCATATGATACCATCG 643  
|||  
Db 461 AATCAGATTCCAAAAAAGTTTACAAAAGGAGATGTGTCAAAGGTTTCTGATGCAACCGTTG 520  
|||  
Qy 644 GTGCTTCCGCTGGGATTTGGACTGTACTATCTCTGATACAAATGCTCTTCAGAGAAAGTTT 703  
|||  
Db 521 GTGCATCTGCTGGGTTTGGTCTACTCTTATCCCTTACAAACAGTAGCCCTTTCGTAAGTTT 580  
|||  
Qy 704 TCACGAAGCATACATCCTCAGCAGTCAACGCACTTGGCCATTTACCAGTCTCTTAGTTGCGA 763  
|||  
Db 581 TAAAGAAGCAAACTTTCTCAGAAAGTTATAAATATGATAATCTACATGAGTCTAGTGSCCA 640  
|||  
Qy 764 GTTGTGTAGTTCTCATAGGACTTTTTCGAAGTGGAGAGTGGGAAACTTTCGCAAGTCAGA 823  
|||  
Db 641 GTTGTGTAGCGTGGTGGGCTTTTGTCTAGTAGCGAGTGGAAAACCTTTTGAGCAGTGAAA 700  
|||  
Qy 824 TGAGAAACTACAACTCGGGAAGTGTATATGTTTTTGTAGCTTTAGCCCTCGGCGAGTATT 883  
|||  
Db 701 TGGAAAACCTACAACTTGGGAAGGTATCTTATGTATGAACCTAGTGTGAGACAGTGTTA 760  
|||  
Qy 884 CCTGGCAAGTCTACACTCTCTGCTTGTGGGATTTGATCTTCGAGTCACTCTCTGTGTCT 943  
|||  
Db 761 CCTGGCAGGTATTTCTCCATCGTTGCACAGCACTGATCTTTCGAGCTTTCTCCTCATTTCT 820  
|||  
Qy 944 CCAATTCATACAGCTGTGGGATTTGCTATAGTTCCAGTTGGGCGAGTATAGTTTTC 1003  
|||  
Db 821 CAAATGCAATAAGCGCTTTTGGGACTTCCCGTGGTCTCTATCTCTGGCTGTCACTATTTC 880  
|||  
Qy 1004 ATGATAGAATGGAGCATCCAAAATCTTCTCCATTTATTTAGCTATCTGGGGCTTCCTTT 1063  
|||  
Db 881 ATGACAAAATGAACGGCTTAAAGGTGATTTCTATGATTTCTAGCTATTTGGGGTTTCGTAT 940  
|||  
Qy 1064 CATTTGCTTATCAGCACTACCTCGACGAAAAGAGTTGAATATAGCCACACAGTCTGT 1123  
|||

Db 941 CCTATGCTACCAACAATATCTTGTGATGAACAACAACTTTGAAGAAAAGTAATGAATAACCAA 1000

QY 1124 TAGGAGA 1130

Db 1001 CAACAGA 1007

## RESULT 7

ADG88190

XX ID ADG88190 standard; cDNA; 1047 BP.

XX AC ADG88190;

XX DT 22-APR-2004 (first entry)

XX A. thaliana RPP4-upregulated pathogen infection-related gene #632.

XX Pathogen infection-related gene; plant; Peronospora parasitica;  
XX defence mechanism; RPP4; pathogen resistance; transgenic plant; oomycete;  
XX fungus; bacterium; virus; nematode; insect; aphid; gene; ss.

XX Arabidopsis thaliana.

XX WO200222675-A2.

XX 21-MAR-2002.

XX 14-SEP-2001; 2001WO-US028506.

XX 15-SEP-2000; 2000US-0232778P.

XX 22-JUN-2001; 2001US-0300183P.

XX (SYGN ) SYNGENTA PARTICIPATIONS AG.

XX (UNYC-) UNIV NORTH CAROLINA.

XX (GLAZ/) GLAZEBROOK J.

XX (WANG/) WANG X.

XX (DANGL/) DANGL J L.

XX (EULG/) EULGEM T.

XX (ZHUT/) ZHU T.

XX Glazebrook J, Wang X, Dangel JL, Eulgem T, Zhu T;

XX WPI; 2002-292409/33.

XX Novel isolated polynucleotide, useful for conveying pathogen resistance  
XX to plants, and for identifying plants infected with a pathogen.

XX Claim 3; SEQ ID NO 632; 605pp; English.

XX The invention relates to 691 Arabidopsis thaliana genes (ADG87559--  
XX ADG87557)) whose expression is altered in response to pathogen infection,  
XX and to homologues of these genes from other plants or fungi, especially  
XX from maize, soybean, barley, alfalfa, sunflower, canola (oilseed rape),  
XX cotton, peanut, sorghum, tobacco, sugarbeet, rice or wheat. The  
XX expression of genes of the invention was upregulated or downregulated in  
XX Arabidopsis plants infected with the oomycete Peronospora parasitica,  
XX indicating that they play a role in defence mechanisms. The genes of the  
XX invention are regulated by RPP7 or RPP8 which act via unconventional  
XX signalling cascades, or by the RPP4-dependent pathway. The invention also  
XX relates to polypeptides encoded by the pathogen infection-related genes;  
XX promoter motifs from pathogen infection-related genes (ADG88243-ADG88327)  
XX; expression cassettes, host cells and pathogen-resistant transgenic  
XX plants and their progeny comprising a polynucleotide of the invention;  
XX and a method of identifying a plant cell infected with a pathogen. The  
XX polynucleotide sequences and methods of the invention are useful for  
XX identifying plants infected with a pathogen, and for conferring  
XX resistance to pathogens such as oomycetes, fungi, bacteria, viruses,  
XX nematodes and insects (e.g., aphids). The present sequence represents an  
XX Arabidopsis thaliana gene whose expression is altered in response to  
XX Peronospora parasitica infection. Note: The sequence data for this patent  
XX did not form part of the printed specification, but was obtained in  
XX electronic format directly from WIPO at  
XX ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 1047 BP; 309 A; 225 C; 199 G; 314 T; 0 U; 0 Other;

Query Match 31.0%; Score 369.6; DB 6; Length 1047;  
Best Local Similarity 62.2%; Pred. No. 1.8e-102;  
Matches 601; Conservative 0; Mismatches 359; Indels 6; Gaps 1;

QY 197 GATTGTACTATGAAATGGAGGAAATAGCACATATGTGTAACACACTTCTTCAACTCATTTG 256

Db 71 GAGTATACTATGACACGGAGGAAACAGTAATGGCTAGCAACGGTAGTTCAACTGTTG 130

QY 257 GCTTCCCTGTACTGGTTCTGTTCCGCTTCTTTCTCGAATCAGGCAACCCAAATCAACAG 316

Db 131 GCTTTCTGTGTACTTCCATATATATCTTGTCTATTA-----AAACACATGCAACAA 184

QY 317 ATACAAATTCAGTCAGTCCCTTCTTCCACACCCCTTGCATCGGTTTACTTGTGCACGT 376

Db 185 CTGATAGAGATGGAAGAAACCTCACCTAGGAACCGTGTATTTGGTTTACGTAGTCTTG 244

QY 377 GACTGTAGTGTCCGCTTATGCTTATTTGTCTGCAGTAGGGTTGCTCTACTTTACCACTCT 436

Db 245 GACTTCTTTGTAGGAGCAGATTGCTATCTGTACTCCATTGGACTTCTTACTTACCCTTT 304

QY 437 CTACTTTCTCCCTCATCTTTGGCCTCACAGTTGGCCTTCACTCCCTTTTCTCATATTTCC 496

Db 305 CTACCTATTTCCCTGATCTGTGCATCTCAGTTAGCCTTCAATGCTTCTTCTCTTATTTTC 364

QY 497 TTAACTCGCAAAAGTTCACTCCTTTGATAGTAGTCAGTTCTTTGGTCTTCTCCACATCT 556

Db 365 TTAACTCACAACAACTTACCCCTATCATTTTAAATTTCTTTCTTCTCTCTTAACTATAT 424

QY 557 CTGCTCTTCTTGTGGTCAACACTGATTCAGAAAACCTCAACTAATGTATCTAGAGTACACT 616

Db 425 CCACCTACTTGCATTCATTAATAGGAGACAGACTCCACAAGAGTTACAAAAGGAGAGT 484

QY 617 ATGTGATCGGGTTTCATATGATACCATCGGTCTTCCGCTGGGATTTGGACTTTACTATCTC 676

Db 485 ATGTCAAAGGTTTCATATGACACCGTTGCTGCTGCTGCTGTTATGGTCTTAGTCTTATCCC 544

QY 677 TGATACAAATGCTCTTTCAGGAAGTTTTCACGAAGCATACATCTTCAGCAGTACAGGACT 736

Db 545 TACAACAGCTAGCCTTTCTAAAGTCTTAAAGAGCAAAATTTCTCAGAAGTTATGGATA 604

QY 737 TGGCCATTTACAGTCTCTAGTTGGAGTTGTGTAGTTCTCATAGGACTTTTTGCAAGTG 796

Db 605 TGATATCTACGTGAGTCTAGTGGCCAGTTGTGTTAGCGTGGTGGGCTTTTGTGTAGCA 664

QY 797 GAGAGTGGGAACTTTTGCCTGAGATGAGAACTACAAACTCGGAAAAGTGTCAATG 856

Db 665 GTGAGTGGGAAACTTTTGAGCAGTGAATGGATAACTACAAACATGGGAAGGTATCCTACA 724

QY 857 TTTTCACTTTAGCCTCGGCAGCTATTTCTTGGCAAGTCTACACTCCTGCTTGTGGGAT 916

Db 725 TTATGAACCTAGTGTGGACAGCTTACTTGGCAGGTATTTCTCCATCGGTGGCACAGGAC 784

QY 917 TGATCTTCGAGTCATCTCTGTTCTCCAAATTCATTAACAGCTGTGGGATTTGCTATAG 976

Db 785 TGATCTTCGAGCTCTCTCTCTATTTCTCAATGCAATGAGCGTTTGGGACTCCAGTGG 844

QY 977 TTCCAGTTTCGGCAGTGTAGTTTTTCCATGATAGAAATGGACGATCCAAAATTTCTTCCA 1036

Db 845 TTCTCTCTTGGCTGTAATCATTTTTTCCATGACAAAATGAATGGGTTAAAGGTTGATTTCTA 904

QY 1037 TTAATTTTACGTATCTGGCGCTTCTTTTCATTCGTTCTATCAGCACTACCTCGACGAAGA 1096

Db 905 TGATCCTAGCTATTTGGGGTTTCACTTCTCTATGCTACCAACAATATCTTGTATGACAAA 964

QY 1097 AGTTGAATACTAGCCACAAAGTGTGTAGGAGATCTTCTATCTACTGTTTGGAGGAGTTC 1156

Db 965 ACTTGAAGAAAATCATGAAATCAACAACAGAAATCCCTTGACCCACCAAGAGCAAG 1024

QY 1157 ACACAA 1162





PF 19-FEB-1999; 99DE-01007209.  
XX PR 19-FEB-1999; 99DE-01007209.  
XX (FROM/) FROMMER W.  
XX PI Gillissen B, Buerkle L, Andre B, Frommer WB;  
XX WPI; 2000-566202/53.  
XX Nucleic acid, useful for producing transgenic plants with altered  
PT nucleobase transport, encodes a nucleobase transporter protein of  
PT Arabidopsis thaliana.  
XX PS Claim 1f; Page 12-13; 24pp; German.  
XX This invention describes a novel nucleic acid encoding a plant nucleobase  
CC transporter (I). (I) is produced by complementation of a nucleobase  
CC transport (NBT)-defective host cell with a plant gene bank by selection  
CC of NBT-positive cells. (I) is used to isolate homologous sequences from  
CC bacteria, fungi, plants, animals and humans, for expression of the  
CC encoded protein (II) in prokaryotic or eukaryotic cells, for inhibiting  
CC expression of (II) (when in antisense orientation), and to produce  
CC transgenic crop plants. The transgenic plants have modified nucleobase  
CC transport properties, e.g. altered affinity and substrate specificity  
CC that may result in more efficient nucleobase transport in leaves, changes  
CC in apical dominance, flowering behaviour and senescence, or improved  
CC distribution of pesticides. This sequence encodes the Arabidopsis  
CC thaliana PUP1 protein which is described in the method of the invention  
XX  
SQ Sequence 1145 BP; 309 A; 234 C; 226 G; 376 T; 0 U; 0 Other;  
Query Match 26.6%; Score 317.2; DB 3; Length 1145;  
Best Local Similarity 57.5%; Pred. No. 2.3e-86;  
Matches 596; Conservative 0; Mismatches 428; Indels 12; Gaps 1;  
QY 75 GGTGGTAACATGATATCATATCATACAGCTGTGCTCAACCGAGAACTATAAAGGTGGCT 134  
DB 18 GGTAGAACCTCGAGGGAAATTTTCAACAGAGAGAGAGTCAACAGTACTCTTGGAGGTT 77  
QY 135 TGGTGTCTCATATAGTATCTTTGCTCTTTTCCAGCCACTAGCTACAAATCTGGG 194  
DB 78 AAGAGTGTCTCTATGTCT 137  
QY 195 TAGATTGTACTATGAAATGAGGAAATAGCAATATGTTGTTAAACACTTCTTCAACTCAT 254  
DB 138 TAGACTTTACTAGAAAGGCGGTAAAGACATAGGCTCGAAACCTGGTTTCACTTGT 197  
QY 255 TGGTTCCTCTACTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCT 314  
DB 198 AGGGTTTCTTTAAACCTTCTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCT 257  
QY 315 AGATACAAATTTTCAGTCAGTCCCTTCTCTCCACCTTGCATCGGTTTACTTGTGCAC 374  
DB 258 AACCATTTACAAATAAATACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 317  
QY 375 TGGACTGTCTAGTGTCTGCTTATGTTCTGCTGAGTGGTGTGCTCTACTTACCACT 434  
DB 318 TGGTGTCTGTTGCTGGACATGATATTTTGTACTCATTTGGGCTACTTACTTCTCTGT 377  
QY 435 CTCTACTTTCTCCCTCATCTTGGCTCTCAAGTGTGGCTTCACTGCTTTTCTCATATTT 494  
DB 378 CTCAACTTCTCTTTGATCTCTGCGTGGCAATTTGGCTTTTAAACGCGCTCTCTCTACTT 437  
QY 495 CCTTAACCTGCAAAAGTTTCACTCTTGTATGATGAGTCTTTTGTCTTCTCTCACTGATC 554  
DB 438 CTTAACTCACAAAAAATCACACCAATTTATCTCAATTCACATGTTCTCTTAAACCATATC 497  
QY 555 CTCTGCTCTTCTTGTGTTCAACACTGATTCAGAA-----AACTCAACTAAGT 602  
DB 498 TTCTACACTTCTTGTATTCACATGAAACCAATCTCCCTCTTCTACTTCAAGTCCGC 557  
QY 603 ATCTAGAGTACAGTATGTGATCGGGTTCATATGTACCATCGGTGCTTCCGCTGGGATTGG 662

DB 558 AGCCAAGTCCAAGTATGTGATGGATACATCTCGCGGTGCGTAGCTCAGCTGGTTATTC 617  
QY 663 ACTGTTACTATCTCTGATACAAATGCTCTTTAGGAAAGTTTTCAGAACATACATCCTC 722  
DB 618 TCTGGTCTTTCTTTTAAACAGATTACGCGTTTGAAGAGATTCTAAAGAAATACACATCAA 677  
QY 723 AGCAGTCACGGACTTGGCCATTTACAGTCTCTAGTTGCGAGTTTGTGTAGTTCTCATAGG 782  
DB 678 GGCTATTTTAGACATGGCCACATATCCGTTCTATGTTAGTACTTGTGTAGTTGTGTAGG 737  
QY 793 ACTTTTTCAGAGTGGAGAGTGGGAACTTTTGCACAGTGGAGTGAAGAACTACAACTCGG 842  
DB 738 ACTTTTTCGAGTGGTGGTGGGAAAGCTGAGTACAGAAATGGAAGAGTTTCAACTAGG 797  
QY 843 GAAAGTGTATATGTTTGTGACTTTTAGCTTCGGCAGCTATTTCTCGCAAGTCTACACTCC 902  
DB 798 GAAAGCTCTATACATTTTGTATTAACATCGTTTCAACGATATCATGGCAAGCTTGTGTAT 857  
QY 903 TGGTCTTGTGGGATTGATCTTTCGAGTCACTCTCTGTTCTTCTTCCAAATTCATTAACAGCTGT 962  
DB 858 TGGAAAGTGTGTTGTTGATTATCGAAGTTTTCATCGCTTTTTCATATGTCATAAGCACTCT 917  
QY 963 GGGATTGCTATAGTTTCCAGTTGCGGCAAGTGAATGTTTCCATGATAGAAATGGAGCATC 1022  
DB 918 TTGTTTACCAAGTTGTGCTGTTTCTTGTGTTGTTCTTCTTCCGTGATGAGTGGGAAT 977  
QY 1023 CAAATCTTCTCCATTTTATGCTATCTGGGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1082  
DB 978 CAAGTTGGTGAATGTTTGTGGCCATCTGGGGATTTGTTCTTATGTTTATCAGCATTTA 1037  
QY 1083 CCTCGACGAAAGAAAG 1098  
DB 1038 TGTCAATGATAGAAAG 1053  
RESULT 10  
AAC51119  
ID AAC51119 standard; DNA; 1292 BP.  
XX AC AAC51119;  
XX DT 18-OCT-2000 (first entry)  
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 67344.  
KW Hybridisation assay; genetic mapping; gene expression control;  
KW protein identification; signal transduction pathway; metabolic pathway;  
KW promoter; termination sequence; ss.  
OS Arabidopsis thaliana.  
PN EP1033405-A2.  
PD 06-SEP-2000.  
XX 25-FEB-2000; 2000EP-00301439.  
XX 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131449P.

PR	30-APR-1999;	99US-0132048P.	PR	21-JUL-1999;	99US-0145086P.
PR	30-APR-1999;	99US-0132407P.	PR	21-JUL-1999;	99US-0145088P.
PR	04-MAY-1999;	99US-0132484P.	PR	22-JUL-1999;	99US-0145085P.
PR	05-MAY-1999;	99US-0132485P.	PR	22-JUL-1999;	99US-0145087P.
PR	06-MAY-1999;	99US-0132486P.	PR	22-JUL-1999;	99US-0145089P.
PR	06-MAY-1999;	99US-0132487P.	PR	22-JUL-1999;	99US-0145192P.
PR	07-MAY-1999;	99US-0132863P.	PR	23-JUL-1999;	99US-0145145P.
PR	11-MAY-1999;	99US-0134256P.	PR	23-JUL-1999;	99US-0145218P.
PR	14-MAY-1999;	99US-0134218P.	PR	23-JUL-1999;	99US-0145224P.
PR	14-MAY-1999;	99US-0134219P.	PR	26-JUL-1999;	99US-0145276P.
PR	14-MAY-1999;	99US-0134221P.	PR	27-JUL-1999;	99US-0145913P.
PR	14-MAY-1999;	99US-0134370P.	PR	27-JUL-1999;	99US-0145918P.
PR	18-MAY-1999;	99US-0134768P.	PR	27-JUL-1999;	99US-0145919P.
PR	19-MAY-1999;	99US-0134941P.	PR	28-JUL-1999;	99US-0145951P.
PR	20-MAY-1999;	99US-0135124P.	PR	02-AUG-1999;	99US-0146386P.
PR	21-MAY-1999;	99US-0135353P.	PR	02-AUG-1999;	99US-0146388P.
PR	24-MAY-1999;	99US-0135629P.	PR	02-AUG-1999;	99US-0146389P.
PR	25-MAY-1999;	99US-0136021P.	PR	03-AUG-1999;	99US-0147038P.
PR	27-MAY-1999;	99US-0136392P.	PR	04-AUG-1999;	99US-0147204P.
PR	28-MAY-1999;	99US-0136782P.	PR	04-AUG-1999;	99US-0147302P.
PR	01-JUN-1999;	99US-0137222P.	PR	05-AUG-1999;	99US-0147302P.
PR	03-JUN-1999;	99US-0137528P.	PR	06-AUG-1999;	99US-0147260P.
PR	04-JUN-1999;	99US-0137502P.	PR	06-AUG-1999;	99US-0147303P.
PR	07-JUN-1999;	99US-0137724P.	PR	06-AUG-1999;	99US-0147416P.
PR	08-JUN-1999;	99US-0138094P.	PR	09-AUG-1999;	99US-0147493P.
PR	10-JUN-1999;	99US-0138540P.	PR	09-AUG-1999;	99US-0147935P.
PR	10-JUN-1999;	99US-0138847P.	PR	10-AUG-1999;	99US-0148171P.
PR	14-JUN-1999;	99US-0139111P.	PR	11-AUG-1999;	99US-0148319P.
PR	16-JUN-1999;	99US-0139452P.	PR	13-AUG-1999;	99US-0148341P.
PR	16-JUN-1999;	99US-0139453P.	PR	13-AUG-1999;	99US-0148565P.
PR	17-JUN-1999;	99US-0139492P.	PR	13-AUG-1999;	99US-0148684P.
PR	18-JUN-1999;	99US-0139454P.	PR	16-AUG-1999;	99US-0149368P.
PR	18-JUN-1999;	99US-0139455P.	PR	17-AUG-1999;	99US-0149175P.
PR	18-JUN-1999;	99US-0139456P.	PR	18-AUG-1999;	99US-0149426P.
PR	18-JUN-1999;	99US-0139457P.	PR	20-AUG-1999;	99US-0149722P.
PR	18-JUN-1999;	99US-0139458P.	PR	20-AUG-1999;	99US-0149723P.
PR	18-JUN-1999;	99US-0139459P.	PR	20-AUG-1999;	99US-0149929P.
PR	18-JUN-1999;	99US-0139460P.	PR	23-AUG-1999;	99US-0149902P.
PR	18-JUN-1999;	99US-0139461P.	PR	23-AUG-1999;	99US-0149930P.
PR	18-JUN-1999;	99US-0139462P.	PR	25-AUG-1999;	99US-0150566P.
PR	18-JUN-1999;	99US-0139463P.	PR	26-AUG-1999;	99US-0150884P.
PR	18-JUN-1999;	99US-0139750P.	PR	27-AUG-1999;	99US-0151065P.
PR	18-JUN-1999;	99US-0139763P.	PR	27-AUG-1999;	99US-0151066P.
PR	21-JUN-1999;	99US-0139817P.	PR	27-AUG-1999;	99US-0151080P.
PR	22-JUN-1999;	99US-0139899P.	PR	30-AUG-1999;	99US-0151303P.
PR	23-JUN-1999;	99US-0140353P.	PR	31-AUG-1999;	99US-0151438P.
PR	23-JUN-1999;	99US-0140354P.	PR	01-SEP-1999;	99US-0151930P.
PR	24-JUN-1999;	99US-0140695P.	PR	07-SEP-1999;	99US-0152363P.
PR	28-JUN-1999;	99US-0140823P.	PR	10-SEP-1999;	99US-0153070P.
PR	29-JUN-1999;	99US-0140991P.	PR	13-SEP-1999;	99US-0153758P.
PR	30-JUN-1999;	99US-0141287P.	PR	13-SEP-1999;	99US-0154018P.
PR	01-JUL-1999;	99US-0141842P.	PR	16-SEP-1999;	99US-0154039P.
PR	01-JUL-1999;	99US-0142154P.	PR	20-SEP-1999;	99US-0154779P.
PR	02-JUL-1999;	99US-0142055P.	PR	22-SEP-1999;	99US-0155139P.
PR	06-JUL-1999;	99US-0142390P.	PR	23-SEP-1999;	99US-0155486P.
PR	08-JUL-1999;	99US-0142803P.	PR	24-SEP-1999;	99US-0155659P.
PR	09-JUL-1999;	99US-0142920P.	PR	28-SEP-1999;	99US-0156458P.
PR	12-JUL-1999;	99US-0142977P.	PR	29-SEP-1999;	99US-0156596P.
PR	13-JUL-1999;	99US-0143542P.	PR	04-OCT-1999;	99US-0157117P.
PR	14-JUL-1999;	99US-0143624P.	PR	05-OCT-1999;	99US-0157753P.
PR	15-JUL-1999;	99US-0144005P.	PR	06-OCT-1999;	99US-0157865P.
PR	16-JUL-1999;	99US-0144085P.	PR	07-OCT-1999;	99US-0158029P.
PR	16-JUL-1999;	99US-0144086P.	PR	08-OCT-1999;	99US-0158232P.
PR	19-JUL-1999;	99US-0144332P.	PR	12-OCT-1999;	99US-0158369P.
PR	19-JUL-1999;	99US-0144331P.	PR	13-OCT-1999;	99US-0159293P.
PR	19-JUL-1999;	99US-0144332P.	PR	13-OCT-1999;	99US-0159294P.
PR	19-JUL-1999;	99US-0144333P.	PR	13-OCT-1999;	99US-0159295P.
PR	19-JUL-1999;	99US-0144334P.	PR	14-OCT-1999;	99US-0159329P.
PR	19-JUL-1999;	99US-0144335P.	PR	14-OCT-1999;	99US-0159330P.
PR	20-JUL-1999;	99US-0144352P.	PR	14-OCT-1999;	99US-0159331P.
PR	20-JUL-1999;	99US-0144632P.	PR	14-OCT-1999;	99US-0159637P.
PR	20-JUL-1999;	99US-0144884P.	PR	14-OCT-1999;	99US-0159638P.
PR	21-JUL-1999;	99US-0144814P.	PR	18-OCT-1999;	99US-0159584P.

PR	21-OCT-1999;	99US-0160741P.	
PR	21-OCT-1999;	99US-0160767P.	
PR	21-OCT-1999;	99US-0160768P.	
PR	21-OCT-1999;	99US-0160770P.	
PR	21-OCT-1999;	99US-0160814P.	
PR	21-OCT-1999;	99US-0160815P.	
PR	22-OCT-1999;	99US-0160980P.	
PR	22-OCT-1999;	99US-0160981P.	
PR	22-OCT-1999;	99US-0160989P.	
PR	25-OCT-1999;	99US-0161404P.	
PR	25-OCT-1999;	99US-0161405P.	
PR	25-OCT-1999;	99US-0161406P.	
PR	26-OCT-1999;	99US-0161359P.	
PR	26-OCT-1999;	99US-0161360P.	
PR	26-OCT-1999;	99US-0161361P.	
PR	28-OCT-1999;	99US-0161920P.	
PR	28-OCT-1999;	99US-0161992P.	
PR	28-OCT-1999;	99US-0161993P.	
PR	29-OCT-1999;	99US-0162142P.	
Query Match			26.5%; Score 316.8; DB 3; Length 1292;
Best Local Similarity			57.1%; Pred. No. 3.2e-86;
Matches 604; Conservative 0; Mismatches 442; Indels 12; Gaps 1;			
QY	53	AAGCAACCTTAGATCATGAGGTGTAACCTGAATCATCATCATCAGCTGTGCCTCAAA	112
Db	47	AGAAGCTCCATCTTACGTAATGCGCAACCTGAGGGAATTTTCAACAGAGAGAGAA	106
QY	113	CGGAGAACTATAAAGTGCCTTCGTCTCCATATACGTAACTTTTGTCTTTTGGC	172
Db	107	GTCAAGTACTCTTGAGGTGAAGAGTGTCTCTATGTCACTCTCTCTTAGCTGGAG	166
QY	173	AGCCACTAGCTACAATTTGGGTAGATTTGTAATGAAATGAGGAAATAGCACATATG	232
Db	167	AGACAATAGCACTCTCTTAGGTAGACTTTACTACGAAAGGCGTAAAGCACATGGC	226
QY	233	TGGTAACACTTTTCAACTATTTGGCTTCCGTGACTGTTGCTGTTGTTCCGCTCTTTTTC	292
Db	227	TCGAAACCTTGGTTGAGCTTGTAGGGTTTCTTTAAACCCCTTCTCTTATTATTAA	286
QY	293	GAATCAGGCAACCCAACTCAACAGATACAAATTTCACTCAGTCCCTTCTTACCACCC	352
Db	287	AGCTGAGCGCTCAAGACTAAACCAATTAACCAAAAAAATACTCTCTCTTCTTGACAC	346
QY	353	TTGCATCGGTTTACTTTGTGACCTGGACTGCTAGTGTCCGCTATGTCTTATTTGTCTGCAG	412
Db	347	TATCTTTAGTGTATATGACTTGGCTTGGCTTGTGCTGGACATTTATTTTGTACTCAT	406
QY	413	TAGGGTTGCTCTACTTACAGTCTCTACTTTCTCCCTCATCTTTGGCTCACAAGTTGGCCT	472
Db	407	TTGGGCTACTTTACCTTCTCTCACTTCTCACTTCTCTTTGATCTCTGGCTCGCAATTGGCTT	466
QY	473	TCACTGCCCTTTTCTCATATTTCTTAACTCGCAAAAGTTCACTCTCTTTGATAGTCAGTT	532
Db	467	TTAACCCGCTTCTCTTACTTCTTAACTCAACAAAAATCAACCAATTTATCTCAAT	526
QY	533	CTTTGCTTCTCTCTCACTGTATCTCTGTCTTCTTGTGTCAACACTGATTCAGAA- - -	588
Db	527	CACCTTGTCTTAAACCATCTTCTTACACTTCTTGTATCCAAATGACCAAGATCTC	586
QY	589	-----AAGCACTAATGATATCTAGAGTACAGATATGATGATCGGGTTTATATGACCA	640
Db	587	CCTCTTCTACTTCAAAAGTCCGAGCCAGTCCCAAGTATGTGATTTGGATACATCTGGCGG	646
QY	641	TCGGTCTTCCGCTGGGATGGACTGTCTTACTATCTCTGATACAAATGCTCTTCAGGAAG	700
Db	647	TCGGTAGCTCAGCTGGTTATCTCTGGTGTCTTCTTTAAACAGATTACGGGTTTCGAAAGA	706
QY	701	TTTTCACGAAGCATACCTCTCAGCAGTCCGAGCTTGGCCATTTACCAAGTCTCTAGTTG	760
Db	707	TTCTAAGAAATACACATTTCAAGGCTATTTTAGACATGGCCACATATCGCTCTATGGTAG	766
QY	761	CGAGTTGTGTAGTTCTCATAGGACTTTTTCGAAGTGGAGAGTGGGAAACTTTTGCCAAAGTG	820

Db	767	CTACTTGTGTAGTGTGGTAGGACTTTTGTGAAGTGGTGGTGGAAAAAGCTGAGTACAG	826
QY	821	AGATGAGAACTACAACTCGGGAAGTGTGCATATATGTTTGTACTTTTAGCCTCGGACGTA	880
Db	827	AAATGGAAGAGTTTCAACTAGGGAAGGCTCATACATTTTGATAAAACATCGGTTCAACGA	886
QY	881	TTTCTCTGCAAGTCTACACTCTCTGTCTTGTGGGATTGATCTTCGAGTCATCTCTGTGT	940
Db	887	TATCATGGCAAGCTTTTGTGATTGGAAGTGTGTTGTTGATTATCGAAGTTTCATCGCTT	946
QY	941	TCTCCAAATTCATAACAGCTGTGGGATTGCTTATAGTTCCAGTTGGCGGAGTATAGTTT	1000
Db	947	TTTCCAATGTCTAAGCACTTTTGTTTACCAGTTGTGCTCTTCTTGTCTTGTCTTCT	1006
QY	1001	TCCATGATAGATGGAAGCATCCAAATCTTCTCCATTTATTTTAGTATCTCGGCTCC	1060
Db	1007	TCCGTGATGAGATGAGTGAATCAAGTTGGTGGCAATGTTTGTGGCAATCTGGGATTTG	1066
QY	1061	TTTCAATTCCTTATCAGCACTACCTCGACGAAAAAGAG	1098
Db	1067	TTTCTTATGTTTATCAGCATTATGTCAATGATAGAAAG	1104
RESULT 11			
AAC37289			
ID	AAC37289 standard; DNA; 1295 BP.		
XX	AAC37289;		
AC	AAC37289;		
XX	17-OCT-2000 (first entry)		
XX	Arabidopsis thaliana DNA fragment SEQ ID NO: 16846.		
DE	Hybridisation assay; genetic mapping; gene expression control;		
KW	protein identification; signal transduction pathway; metabolic pathway;		
KW	promoter; termination sequence; ss.		
XX	Arabidopsis thaliana.		
OS	Arabidopsis thaliana.		
PN	EP1033405-A2.		
XX	06-SEP-2000.		
PD	25-FEB-2000; 2000EP-00301439.		
PF	25-FEB-1999; 99US-0121825P.		
PR	05-MAR-1999; 99US-0123180P.		
PR	09-MAR-1999; 99US-0123548P.		
PR	23-MAR-1999; 99US-0125788P.		
PR	25-MAR-1999; 99US-0126264P.		
PR	29-MAR-1999; 99US-0126785P.		
PR	01-APR-1999; 99US-0127462P.		
PR	06-APR-1999; 99US-0128234P.		
PR	08-APR-1999; 99US-0128714P.		
PR	16-APR-1999; 99US-0129845P.		
PR	19-APR-1999; 99US-0130077P.		
PR	21-APR-1999; 99US-0130449P.		
PR	23-APR-1999; 99US-0130510P.		
PR	23-APR-1999; 99US-0130891P.		
PR	28-APR-1999; 99US-0131449P.		
PR	30-APR-1999; 99US-0132048P.		
PR	30-APR-1999; 99US-0132407P.		
PR	04-MAY-1999; 99US-0132484P.		
PR	05-MAY-1999; 99US-0132485P.		
PR	06-MAY-1999; 99US-0132486P.		
PR	07-MAY-1999; 99US-0132487P.		
PR	07-MAY-1999; 99US-0132863P.		
PR	11-MAY-1999; 99US-0134256P.		
PR	14-MAY-1999; 99US-0134218P.		
PR	14-MAY-1999; 99US-0134219P.		
PR	14-MAY-1999; 99US-0134221P.		
PR	14-MAY-1999; 99US-0134370P.		

PR 18-MAY-1999;	99US-0134768P.	PR 27-JUL-1999;	99US-0145919P.
PR 19-MAY-1999;	99US-0134941P.	PR 28-JUL-1999;	99US-0145951P.
PR 20-MAY-1999;	99US-0135124P.	PR 02-AUG-1999;	99US-0146386P.
PR 21-MAY-1999;	99US-0135353P.	PR 02-AUG-1999;	99US-0146388P.
PR 24-MAY-1999;	99US-0135629P.	PR 02-AUG-1999;	99US-0146389P.
PR 25-MAY-1999;	99US-0136021P.	PR 03-AUG-1999;	99US-0147038P.
PR 27-MAY-1999;	99US-0136392P.	PR 04-AUG-1999;	99US-0147204P.
PR 28-MAY-1999;	99US-0136782P.	PR 04-AUG-1999;	99US-0147302P.
PR 01-JUN-1999;	99US-0137222P.	PR 05-AUG-1999;	99US-0147192P.
PR 03-JUN-1999;	99US-0137528P.	PR 05-AUG-1999;	99US-0147260P.
PR 04-JUN-1999;	99US-0137502P.	PR 06-AUG-1999;	99US-0147303P.
PR 07-JUN-1999;	99US-0137724P.	PR 06-AUG-1999;	99US-0147416P.
PR 08-JUN-1999;	99US-0138094P.	PR 09-AUG-1999;	99US-0147493P.
PR 10-JUN-1999;	99US-0138540P.	PR 09-AUG-1999;	99US-0147935P.
PR 10-JUN-1999;	99US-0138847P.	PR 10-AUG-1999;	99US-0148171P.
PR 14-JUN-1999;	99US-0139119P.	PR 11-AUG-1999;	99US-0148319P.
PR 16-JUN-1999;	99US-0139452P.	PR 12-AUG-1999;	99US-0148341P.
PR 16-JUN-1999;	99US-0139453P.	PR 13-AUG-1999;	99US-0148565P.
PR 17-JUN-1999;	99US-0139492P.	PR 13-AUG-1999;	99US-0148684P.
PR 18-JUN-1999;	99US-0139454P.	PR 16-AUG-1999;	99US-0149368P.
PR 18-JUN-1999;	99US-0139455P.	PR 17-AUG-1999;	99US-0149175P.
PR 18-JUN-1999;	99US-0139456P.	PR 18-AUG-1999;	99US-0149426P.
PR 18-JUN-1999;	99US-0139457P.	PR 20-AUG-1999;	99US-0149722P.
PR 18-JUN-1999;	99US-0139458P.	PR 20-AUG-1999;	99US-0149723P.
PR 18-JUN-1999;	99US-0139459P.	PR 20-AUG-1999;	99US-0149929P.
PR 18-JUN-1999;	99US-0139460P.	PR 23-AUG-1999;	99US-0149902P.
PR 18-JUN-1999;	99US-0139461P.	PR 23-AUG-1999;	99US-0149930P.
PR 18-JUN-1999;	99US-0139462P.	PR 25-AUG-1999;	99US-0150566P.
PR 18-JUN-1999;	99US-0139463P.	PR 26-AUG-1999;	99US-0150884P.
PR 18-JUN-1999;	99US-0139750P.	PR 27-AUG-1999;	99US-0151065P.
PR 18-JUN-1999;	99US-0139763P.	PR 27-AUG-1999;	99US-0151066P.
PR 21-JUN-1999;	99US-0139817P.	PR 27-AUG-1999;	99US-0151080P.
PR 22-JUN-1999;	99US-0139899P.	PR 30-AUG-1999;	99US-0151303P.
PR 23-JUN-1999;	99US-0140353P.	PR 31-AUG-1999;	99US-0151438P.
PR 23-JUN-1999;	99US-0140354P.	PR 01-SEP-1999;	99US-0151930P.
PR 24-JUN-1999;	99US-0140695P.	PR 07-SEP-1999;	99US-0152363P.
PR 28-JUN-1999;	99US-0140821P.	PR 10-SEP-1999;	99US-0153070P.
PR 29-JUN-1999;	99US-0140991P.	PR 13-SEP-1999;	99US-0153758P.
PR 30-JUN-1999;	99US-0141287P.	PR 15-SEP-1999;	99US-0154018P.
PR 01-JUL-1999;	99US-0141842P.	PR 16-SEP-1999;	99US-0154039P.
PR 01-JUL-1999;	99US-0142154P.	PR 20-SEP-1999;	99US-0154779P.
PR 02-JUL-1999;	99US-0142055P.	PR 22-SEP-1999;	99US-0155139P.
PR 06-JUL-1999;	99US-0142390P.	PR 23-SEP-1999;	99US-0155486P.
PR 08-JUL-1999;	99US-0142803P.	PR 24-SEP-1999;	99US-0155659P.
PR 09-JUL-1999;	99US-0142920P.	PR 28-SEP-1999;	99US-0156458P.
PR 12-JUL-1999;	99US-0142977P.	PR 29-SEP-1999;	99US-0156596P.
PR 13-JUL-1999;	99US-0143542P.	PR 04-OCT-1999;	99US-0157117P.
PR 14-JUL-1999;	99US-0143624P.	PR 05-OCT-1999;	99US-0157753P.
PR 15-JUL-1999;	99US-0144005P.	PR 06-OCT-1999;	99US-0157865P.
PR 16-JUL-1999;	99US-0144085P.	PR 07-OCT-1999;	99US-0158029P.
PR 16-JUL-1999;	99US-0144086P.	PR 08-OCT-1999;	99US-0158232P.
PR 19-JUL-1999;	99US-0144325P.	PR 12-OCT-1999;	99US-0158369P.
PR 19-JUL-1999;	99US-0144331P.	PR 13-OCT-1999;	99US-0159293P.
PR 19-JUL-1999;	99US-0144332P.	PR 13-OCT-1999;	99US-0159294P.
PR 19-JUL-1999;	99US-0144333P.	PR 13-OCT-1999;	99US-0159295P.
PR 19-JUL-1999;	99US-0144334P.	PR 14-OCT-1999;	99US-0159329P.
PR 19-JUL-1999;	99US-0144335P.	PR 14-OCT-1999;	99US-0159330P.
PR 20-JUL-1999;	99US-0144352P.	PR 14-OCT-1999;	99US-0159331P.
PR 20-JUL-1999;	99US-0144632P.	PR 14-OCT-1999;	99US-0159637P.
PR 20-JUL-1999;	99US-0144884P.	PR 14-OCT-1999;	99US-0159638P.
PR 21-JUL-1999;	99US-0144814P.	PR 18-OCT-1999;	99US-0159584P.
PR 21-JUL-1999;	99US-0145086P.	PR 21-OCT-1999;	99US-0160741P.
PR 21-JUL-1999;	99US-0145088P.	PR 21-OCT-1999;	99US-0160767P.
PR 22-JUL-1999;	99US-0145085P.	PR 21-OCT-1999;	99US-0160768P.
PR 22-JUL-1999;	99US-0145087P.	PR 21-OCT-1999;	99US-0160770P.
PR 22-JUL-1999;	99US-0145089P.	PR 21-OCT-1999;	99US-0160814P.
PR 23-JUL-1999;	99US-0145192P.	PR 21-OCT-1999;	99US-0160815P.
PR 23-JUL-1999;	99US-0145145P.	PR 22-OCT-1999;	99US-0160980P.
PR 23-JUL-1999;	99US-0145218P.	PR 22-OCT-1999;	99US-0160981P.
PR 23-JUL-1999;	99US-0145221P.	PR 22-OCT-1999;	99US-0160989P.
PR 26-JUL-1999;	99US-0145276P.	PR 25-OCT-1999;	99US-0161404P.
PR 27-JUL-1999;	99US-0145913P.	PR 25-OCT-1999;	99US-0161405P.
PR 27-JUL-1999;	99US-0145918P.	PR 25-OCT-1999;	99US-0161406P.

PR	26-OCT-1999;	99US-01611359P.	
PR	26-OCT-1999;	99US-01611360P.	
PR	26-OCT-1999;	99US-01611361P.	
PR	28-OCT-1999;	99US-01611920P.	
PR	28-OCT-1999;	99US-01611922P.	
PR	28-OCT-1999;	99US-01611933P.	
PR	28-OCT-1999;	99US-0162142P.	
Query Match			
Best Local Similarity 57.1%; Score 316.8; DB 3; Length 1295;			
Matches 604; Conservative 0; Mismatches 442; Indels 12; Gaps 1;			
QY	53	AAGCAAACTTATAGATCATGAGTGGTAAGTAACTGAAATCATCATCATGCTGGCTCCAAA	112
Db	48	AAGAACTCCATCTTACGTAATAGGGAACCTGAAGGGAAATTTTCAACAGAAGAGAGAA	107
QY	113	CCGAGAACTATAAAAGTGGCTTCGTGCTGCCATATACGTAATCTTTTTCCTCTTTTGGC	172
Db	108	GTCAAGAGTACTCTTGAGAGTTAAGAGTGTCTCTATGTCACTCTCTCTTAGCTGGAG	167
QY	173	AGCCACTAGCTACAATTTCTGGTGTAGATGTACTATGAAATGGAGGAAATAGCACATATG	232
Db	168	AGCAATAGCCACTCTCTTAGGTAGACTTTTACTACGAAAGGCGTAAAGCACATGGC	227
QY	233	TGTTAACACTTTCTTAACATATGGGCTTCCTGTACTGGTGTCTGTTTCGCGCTCTTTTCTC	292
Db	228	TCGAAACCTTGGTTACGTTGTAGGGTTTCCTTTAAACCTCTCCTTCTATTAATTAATAA	287
QY	293	GAATCAGGAAACCAATCAACAGATACAAATTTTCAGTCAGTCCCTTCTTCCACACCC	352
Db	288	AGCTGAGCGGTCCAAGACTAAACCAATTAACCAAAAAAATACTTCTCTCTTGTGACAC	347
QY	353	TTGCATCGGTTTACTTTGTGCACGTGACTGTAGTGTCCGCTTATGTCTTATTTGTCTGCAG	412
Db	348	TATCTTTAGTGATATTTGACTTGGCTTGTGCTGGACATGTATTTTGTACTCAT	407
QY	413	TAGGGTGTCTACTTACAGTCTCTACTTTCTCCCTCATCTTGGCTCACAGTGGCT	472
Db	408	TTGGGTACTTTTACCTTCTGTCTCAACTTCTCTTTGTATCTCTGGTGCCTAATGGCTT	467
QY	473	TCAGTCCCTTTTCTCATATTTCTTAATCCGAAAGTTTCACTCTTGTGATAGTCAGTT	532
Db	468	TTAACCCGCTTCTCTTACTTCTTAATCTCAAAAAATCAACCAATTTTACTCAAT	527
QY	533	CTTTGTCTCTCTCACTGTATCTCTGTCTTCTTGTGCTCAACACTGATTCAGAA- - -	588
Db	528	CACCTGTCTCTTAAACATATCTTCTACACTTCTTGTATCCAAATGACCAAGATCTC	587
QY	589	- - - - - AACTCAACTAATGATCTAGAGTACAGTATGATGATCGGGTTTCATATGTACCA	640
Db	588	CCTCTTCTACTTCAAAAGTCCGACGCCAAGTCCAAGTATGTGATGGATACATCTGCGGG	647
QY	641	TCGGTCTTCCGCTGGATGAGTGTATCTATCTCTGATACAAATGCTCTTCAGAAAG	700
Db	648	TCGGTAGCTCAGCTGGTTATCTCTGGTGTCTTCTTTAAACAGATTACGGGTTTCGAAAGA	707
QY	701	TTTTCACGAAGATACATCCTCAGCAGTCCAGGCTTGGCCATTTACCAAGTCTCTAGTTG	760
Db	708	TTCTAAGAAATACACATTTCAAGGCTATTTTAGACATGGCCACATATCCGCTATGGTAG	767
QY	761	CGAGTTGTGTAGTTCTATAGGACTTTTTCGAAGTGGAGAGTGGGAAACTTTGCCAAGTG	820
Db	768	CTACTGTGTAGTTGTGGTAGGACTTTTGTGAAGTGGTGGGTGGAAAGGCTGAGTACAG	827
QY	821	AGATGAAACTACAACTCGGAAAGTGTCAATATGTTTGTAGCTTTAGCTCGGAGCTA	880
Db	828	AAATGGAAGAGTTTCAACTAGGGAAGAGCTCTACATTTTGTATAAATCATCGGTTCAACGA	887
QY	881	TTTCTCTGGAAGTCTACACTCTCTGCTTGTGCGGATTTGATCTTCAGTCATCTCTGTGT	940
Db	888	TATCATGGCAAGCTTGTGTGATGGAGTGTGGTTGGTTGATATCGAAGTTTCATCGCTTT	947
QY	941	TCTCCAATTCATAACAGCTGTGGATTGCTTATAGTTCCAGTTGCGGAGTGATAGTTT	1000

Db	948	TTTCCAATGTCATAAGCACTCTTTGTTTACCAGTTGTGCTGTCTTGTCTGTCTTCT 1007
QY	1001	TCCATGATAGAAATGGAGCATCCAAAATCTTCTCCATTTATTTAGCTATCTCGGCTTCC 1060
Db	1008	TCCGTATGAGATGAGTGGATCAAGTTGATTTGCAATGTTTGGCCATCTGGGATTTG 1067
QY	1061	TTTCAATTCGTCTATCAGCACTACCTCGACGAAAGAAAG 1098
Db	1068	TTTCTTATGTTTATCAGCATTATGTCATGATAGAAAG 1105
RESULT 12		
AAC36845		
ID	AAC36845	standard; DNA; 492 BP.
XX	AAC36845;	
XX	AC	
XX	17-OCT-2000	(first entry)
XX	Arabidopsis thaliana	DNA fragment SEQ ID NO: 15272.
DE	Hybridisation assay; genetic mapping; gene expression control;	
KW	protein identification; signal transduction pathway; metabolic pathway;	
KW	promoter; termination sequence; ss.	
XX	Arabidopsis thaliana.	
XX	EP1033405-A2.	
XX	06-SEP-2000.	
XX	25-FEB-2000;	2000EP-00301439.
XX	25-FEB-1999;	99US-0121825P.
PR	05-MAR-1999;	99US-0123180P.
PR	09-MAR-1999;	99US-0123548P.
PR	23-MAR-1999;	99US-0125788P.
PR	25-MAR-1999;	99US-0126264P.
PR	29-MAR-1999;	99US-0126785P.
PR	01-APR-1999;	99US-0127462P.
PR	06-APR-1999;	99US-0128234P.
PR	08-APR-1999;	99US-0128714P.
PR	16-APR-1999;	99US-0129845P.
PR	19-APR-1999;	99US-0130077P.
PR	21-APR-1999;	99US-0130449P.
PR	23-APR-1999;	99US-0130510P.
PR	23-APR-1999;	99US-0130891P.
PR	28-APR-1999;	99US-0131449P.
PR	30-APR-1999;	99US-0132048P.
PR	04-MAY-1999;	99US-0132407P.
PR	05-MAY-1999;	99US-0132484P.
PR	06-MAY-1999;	99US-0132485P.
PR	06-MAY-1999;	99US-0132486P.
PR	07-MAY-1999;	99US-0132487P.
PR	07-MAY-1999;	99US-0132863P.
PR	11-MAY-1999;	99US-0134256P.
PR	14-MAY-1999;	99US-0134218P.
PR	14-MAY-1999;	99US-0134219P.
PR	14-MAY-1999;	99US-0134221P.
PR	18-MAY-1999;	99US-0134370P.
PR	19-MAY-1999;	99US-0134768P.
PR	20-MAY-1999;	99US-0134941P.
PR	21-MAY-1999;	99US-0135124P.
PR	21-MAY-1999;	99US-0135353P.
PR	24-MAY-1999;	99US-0135629P.
PR	25-MAY-1999;	99US-0136021P.
PR	27-MAY-1999;	99US-0136392P.
PR	28-MAY-1999;	99US-0136782P.
PR	01-JUN-1999;	99US-0137222P.
PR	03-JUN-1999;	99US-0137528P.
PR	04-JUN-1999;	99US-0137502P.
PR	07-JUN-1999;	99US-0137724P.

PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.

PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 18-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161922P.  
PR 28-OCT-1999; 99US-0161933P.  
PR 29-OCT-1999; 99US-0162142P.

Query Match 25.6%; Score 305.8; DB 3; Length 492;  
Best Local Similarity 90.0%; Pred. No. 4.8e-83;  
Matches 341; Conservative 1; Mismatches 28; Indels 9; Gaps 1;

QY	37	GGTGACCAGAACTTAGAGCAAAACCTTATAGATCATGAGGTGTAACCTGAATCATCATCA	96	PR	11-MAY-1999;	99US-0134256P.
Db	123	GGTGACCAGAACTTAGAGCAAAACCTTCTAGATCATGAG-----GAAACTGAATCA	173	PR	14-MAY-1999;	99US-0134218P.
QY	97	TCAGCTGTGCTCTCAAAACCGAGAAGCTATAAAAGGTGGCTTCGTCTCTCCATATACGTAATC	156	PR	14-MAY-1999;	99US-0134221P.
Db	174	TTTTTCAGTACTCAAAACGAGAAGCTGTAAAGAGGTGGCTCGGTGTCTCCATATACGCAATC	233	PR	18-MAY-1999;	99US-0134370P.
QY	157	TTTGTCTCTTTTGGCCAGCCACTAGCTACAAATTCGGGTAGATTGTACTATGAAAAATGGA	216	PR	19-MAY-1999;	99US-0134768P.
Db	234	TTTGTCTATCTCTGCCAACCACTTGCTACAGTTCTGGGTAGACTGTACTATGAAAAATGGA	293	PR	20-MAY-1999;	99US-0134941P.
QY	217	GGAAATAGACATATGTGTTAAACACTCTTCAACTCATTTGGCTTCCTGCTACTGTTCTG	276	PR	21-MAY-1999;	99US-0135124P.
Db	294	GGGAAAGCACATATGTGGTAACACTTCTTCAACTCATTTGGCTTCCTGCTACTGATCTC	353	PR	24-MAY-1999;	99US-0135353P.
QY	277	TTCCGCTCTCTTTCTCGAATCAGGCAACCCAAATCAACAGATACAAATTTTCAGTCAGTCC	336	PR	25-MAY-1999;	99US-0136021P.
Db	354	TTCCGCTCTCTTTCTCGAATCAGGCAACCCAAATCAACAGATACAAATTTTCAGTCAGTCC	413	PR	27-MAY-1999;	99US-0136392P.
QY	337	CTTCTCTTCAACACCTTGCAATCGGTTTACTTGTGCACTGGACTGCTAGTCCGCTTAT	396	PR	28-MAY-1999;	99US-0136782P.
Db	414	CTTCTCTTCAACACCTTGCAATCGGTTTACTTGTGCACTGGACTGCTAGTCCGCTTAT	473	PR	01-JUN-1999;	99US-0137222P.
QY	397	GCTTATTGTCTGCAGTAG	415	PR	04-JUN-1999;	99US-0137502P.
Db	474	GCTTATTGTCTGCAGTTG	492	PR	07-JUN-1999;	99US-0137724P.
RESULT 13						99US-0138094P.
ARAC47974						99US-0138540P.
ID	AAC47974 standard; DNA; 2660 BP.					99US-0138847P.
AC	AAC47974;					99US-0139119P.
XX	18-OCT-2000 (first entry)					99US-0139452P.
DT	Arabidopsis thaliana DNA fragment SEQ ID NO: 55799.					99US-0139453P.
XX	Hybridisation assay; genetic mapping; gene expression control;					99US-0139462P.
KW	protein identification; signal transduction pathway; metabolic pathway;					99US-0139463P.
KW	promoter; termination sequence; ss.					99US-0139750P.
XX	Arabidopsis thaliana.					99US-0139763P.
OS	EP1033405-A2.					99US-0139817P.
XX	06-SEP-2000.					99US-0139899P.
PN	25-FEB-2000; 2000BP-00301439.					99US-0140354P.
PF	25-FEB-1999;					99US-0140695P.
XX	05-MAR-1999;					99US-0140823P.
PR	09-MAR-1999;					99US-0140911P.
PR	23-MAR-1999;					99US-0141287P.
PR	29-MAR-1999;					99US-0141542P.
PR	01-APR-1999;					99US-0142055P.
PR	06-APR-1999;					99US-0142390P.
PR	08-APR-1999;					99US-0142803P.
PR	16-APR-1999;					99US-0142920P.
PR	19-APR-1999;					99US-0142977P.
PR	21-APR-1999;					99US-0143542P.
PR	23-APR-1999;					99US-0143624P.
PR	28-APR-1999;					99US-0144005P.
PR	30-APR-1999;					99US-0144085P.
PR	04-MAY-1999;					99US-0144325P.
PR	05-MAY-1999;					99US-0144331P.
PR	06-MAY-1999;					99US-0144332P.
PR	07-MAY-1999;					99US-0144333P.
PR						99US-0144334P.
PR						99US-0144335P.
PR						99US-0144336P.
PR						99US-0144632P.
PR						99US-0144884P.
PR						99US-0144814P.
PR						99US-0145086P.
PR						99US-0145087P.
PR						99US-0145088P.
PR						99US-0145089P.
PR						99US-0145192P.
PR						99US-0145145P.

PR	23-JUL-1999;	99US-0145218P.
PR	23-JUL-1999;	99US-01452224P.
PR	26-JUL-1999;	99US-0145276P.
PR	27-JUL-1999;	99US-0145913P.
PR	27-JUL-1999;	99US-0145918P.
PR	27-JUL-1999;	99US-0145919P.
PR	28-JUL-1999;	99US-0145951P.
PR	02-AUG-1999;	99US-0146386P.
PR	02-AUG-1999;	99US-0146388P.
PR	02-AUG-1999;	99US-0146389P.
PR	03-AUG-1999;	99US-0147038P.
PR	04-AUG-1999;	99US-0147204P.
PR	04-AUG-1999;	99US-0147302P.
PR	05-AUG-1999;	99US-0147192P.
PR	05-AUG-1999;	99US-0147260P.
PR	06-AUG-1999;	99US-0147303P.
PR	06-AUG-1999;	99US-0147416P.
PR	09-AUG-1999;	99US-0147493P.
PR	10-AUG-1999;	99US-0147935P.
PR	11-AUG-1999;	99US-0148171P.
PR	11-AUG-1999;	99US-0148319P.
PR	12-AUG-1999;	99US-0148341P.
PR	13-AUG-1999;	99US-0148565P.
PR	13-AUG-1999;	99US-0148684P.
PR	16-AUG-1999;	99US-0149368P.
PR	17-AUG-1999;	99US-0149175P.
PR	18-AUG-1999;	99US-0149426P.
PR	20-AUG-1999;	99US-0149722P.
PR	20-AUG-1999;	99US-0149723P.
PR	20-AUG-1999;	99US-0149929P.
PR	23-AUG-1999;	99US-0149902P.
PR	23-AUG-1999;	99US-0149930P.
PR	25-AUG-1999;	99US-0150566P.
PR	26-AUG-1999;	99US-0150884P.
PR	27-AUG-1999;	99US-0151065P.
PR	27-AUG-1999;	99US-0151066P.
PR	27-AUG-1999;	99US-0151080P.
PR	30-AUG-1999;	99US-0151303P.
PR	31-AUG-1999;	99US-0151438P.
PR	01-SEP-1999;	99US-0151930P.
PR	01-SEP-1999;	99US-0152363P.
PR	10-SEP-1999;	99US-0153070P.
PR	13-SEP-1999;	99US-0153758P.
PR	15-SEP-1999;	99US-0154018P.
PR	16-SEP-1999;	99US-0154039P.
PR	20-SEP-1999;	99US-0154779P.
PR	22-SEP-1999;	99US-0155139P.
PR	23-SEP-1999;	99US-0155486P.
PR	24-SEP-1999;	99US-0155659P.
PR	28-SEP-1999;	99US-0156458P.
PR	29-SEP-1999;	99US-0156596P.
PR	04-OCT-1999;	99US-0157117P.
PR	06-OCT-1999;	99US-015753P.
PR	06-OCT-1999;	99US-0157865P.
PR	07-OCT-1999;	99US-0158029P.
PR	08-OCT-1999;	99US-0158232P.
PR	12-OCT-1999;	99US-0158369P.
PR	13-OCT-1999;	99US-0159293P.
PR	13-OCT-1999;	99US-0159294P.
PR	13-OCT-1999;	99US-0159295P.
PR	14-OCT-1999;	99US-0159329P.
PR	14-OCT-1999;	99US-0159330P.
PR	14-OCT-1999;	99US-0159331P.
PR	14-OCT-1999;	99US-0159637P.
PR	18-OCT-1999;	99US-0159638P.
PR	21-OCT-1999;	99US-0159584P.
PR	21-OCT-1999;	99US-0160741P.
PR	21-OCT-1999;	99US-0160767P.
PR	21-OCT-1999;	99US-0160768P.
PR	21-OCT-1999;	99US-0160770P.
PR	21-OCT-1999;	99US-0160814P.
PR	21-OCT-1999;	99US-0160815P.
PR	22-OCT-1999;	99US-0160980P.

PR	22-OCT-1999;	99US-0160981P.
PR	22-OCT-1999;	99US-0160989P.
PR	25-OCT-1999;	99US-0161404P.
PR	25-OCT-1999;	99US-0161405P.
PR	25-OCT-1999;	99US-0161406P.
PR	26-OCT-1999;	99US-0161359P.
PR	26-OCT-1999;	99US-0161360P.
PR	26-OCT-1999;	99US-0161361P.
PR	28-OCT-1999;	99US-0161320P.
PR	28-OCT-1999;	99US-0161992P.
PR	28-OCT-1999;	99US-0161993P.
PR	29-OCT-1999;	99US-0162142P.

Query Match

22.6%; Score 270; DB 3; Length 2660;

Best Local Similarity 55.6%; Pred. No. 1.1e-71;

Matches 544; Conservative 0; Mismatches 425; Indels 9; Gaps 1;

Qy	125	AAAGGTGGCTTCGTCTCCATATACGTAATCTTTGTGCTCTTTTGCAGCCACTAGCTA	184
Db	189	AATGGTGGTTCCTTGATCCGTTAACATCTTTTCTTATTTGGTCAAGCTGCTTCGG	248
Qy	185	CAATTCTGGTAGATTGTACTATGAAATGGAGAAATAGACATATGTGGTAACATTC	244
Db	249	TTCTTCTTGGTAGGTTTACTATGTAAGGTGGAAACAGTAATAATGATGCCAACTCTTG	308
Qy	245	TTCAACTCATTTGGCTTCCCTGTACTGGTTCCTGTTCCGCTTCTTTTCTCGAATCAGGCAAC	304
Db	309	TTCAACGGCTGCTTTTCTTATACTCTATATCCCGCTTTTGGTACTTCCGCTTCGG---	365
Qy	305	CCAAATCAACAGATACAAATTTTCAGTCAGTCCCTTTCCTTCACCACCCTTTGCATCGGTTT	364
Db	366	-----CTAGTGTAGAGTCTTCGGAGAGTTCGGTTCACCTCAAGTACATTTGTTGATCT	419
Qy	365	ACTTGTGCACTGGACTGTAGTGTCCGCTTATGCTTATTTGTCTGCGAGTAGGGTGTCTCT	424
Db	420	ATGTTTTTGTCTCGGTGTGATCATTTGCTGGAGATAATATGTTATATCTCTGTTGGACTTTGT	479
Qy	425	ACTTACAGTCTACTTTCTCCCTCATCTTGGCCCTCACAGTTGGCCTTCACCTGCTCTTT	484
Db	480	ACCTCTCTGCATCGAGGTATTCGCTCATTTTGGCGTACTCAGTTAGCTTTTCAACGGGTGT	539
Qy	485	TCCTATATTTCTTACTCGCAAAAGTTCACTCCTTTTGATAGTCAGTTCTTTGCTTCTCC	544
Db	540	TCCTTATTTTCATCAATGCTCAGAGTTCACTGCTTTTGATTTCTCACTCCGTTGCTCT	599
Qy	545	TCACTGTATCCTCTGCTCTTCTTGTGGTCAACACTGATTCAGAAAACTCAACTAATGAT	604
Db	600	TGTCGTTCTCCGCTGCTTTGATAGTCTCTCAATGATGCGGATACTCTCTTCTGGTGTCT	659
Qy	605	CTAGAGTACAGTATGTATCGGGTTCATATGTACCATCGGTGCTTCGCTGGGATTGGAC	664
Db	660	CCAGGTCTAAGTATATTTGGGGTTTGTGTGTACATTCGCTGCTGTCTCTATTTCTC	719
Qy	665	TGTTACTATCTCTGATACAAATGCTCTTTCAGGAAAGTTTTTCACGAAGCATACATCCTCAG	724
Db	720	TGTTGCTATCTCTTATGCACTTCTCGTTCGAGAAGATTCTGAAGAGGGAGACGTTTTCTG	779
Qy	725	CAGTCACGACTTGGCCATTTCAGGTCCTAGTGTGGAGTTGTGTAGTTCTCATAGGAC	784
Db	780	TGGTTCTTGAATTCGAAATCTACACTTCTTTAGTGGCGACTTGTGTTTCGGTTATAGGC	839
Qy	785	TTTTTCAAGTGGAGAGTGGGAACTTTGCAAGTGAAGATGAGAAATACAACTCGGGA	844
Db	840	TTTTTGTAGCGGGAATGGAAACGCTGCTGGGAAATGGAGGTTATCATTAAGGGC	899
Qy	845	AAGTGTATATGTTTTGACTTTTAGCCTCGGAGCTATTTCTTGGCAAGTCTACACTCCTG	904
Db	900	AAGCCTTATGTACTGACCTTGTCTGGACAGCAGTTACTTGGCAAGTGTGTTCTCTTG	959
Qy	905	GTCTTGTGGGATTGATCTTCGAGTCATCCTCTGTGTCTTCCAATTCATTAACAGCTGTGG	964
Db	960	GAGTCGTGGGTTTGATATTTCTGGTGACGCTGCTTCTCTCAACGCTCATTAGTACGCTCT	1019





KW Plant; bacterial infection; fungal infection; viral infection; rice;  
KW gene; ds.  
XX  
OS Oryza sativa.  
XX  
PN WO2003000898-A1.  
XX  
XX  
PD 03-JAN-2003.  
XX  
XX  
PF 22-JUN-2001; 2001WO-IB001105.  
XX  
XX  
PR 22-JUN-2001; 2001WO-IB001105.  
XX  
XX  
PA (SYGN ) SYNGENTA PARTICIPATIONS AG.  
XX  
PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;  
PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;  
XX  
XX  
DR WPI; 2003-175290/17.  
XX  
PT Identifying at least one gene involved in plant resistance or response to  
PT pathogenic infection for conferring resistance or tolerance to a plant to  
PT bacterial, fungal or viral infection by determining or detecting plant  
PT gene expression.  
XX  
PS Claim 6; SEQ ID NO 4071; 899pp; English.  
XX  
CC The present invention relates to a method (M1) for identifying genes  
CC involved in plant resistance or response to pathogenic infection. M1  
CC comprises identifying a gene whose expression is significantly altered in  
CC the incompatible interaction of plant gene expression relative to  
CC expression of the gene in an uninfected plant in a mutant plant that  
CC does not express a gene associated with response to pathogenic infection,  
CC or in a corresponding incompatible or compatible interaction. (M1) is  
CC useful for conferring resistance to resistance or tolerance to a plant to  
CC bacterial, fungal or viral infection. The present sequence was used to  
CC illustrate the invention.  
XX  
XX  
SQ Sequence 2175 BP; 363 A; 701 C; 671 G; 439 T; 0 U; 1 Other;  
  
Query Match 17.9%; Score 213.2; DB 8; Length 2175;  
Best Local Similarity 52.9%; Pred. No. 2.8e-54;  
Matches 492; Conservative 0; Mismatches 423; Indels 15; Gaps 1;  
  
QY 89 CATCATCATAGCTGTGCTCAACCCGAGACTATAAAGGTGGCTTCGTCTCCATAT 148  
DB 98 CAGCAGCTCCGGAAACACGACGACCAAGCCATTTCGATGGGCGGATGGTGCAGTGG 157  
  
QY 149 ACCTAATCTTTTGTCTCTTTTCCAGCCACTAGCTACAATTCCTGGTAGATTGTACTATG 208  
DB 158 ACGTCTTCTTCTCATCCCGGTGAGACATCGGCGACACTGCTGGCGAGATACTACTACA 217  
  
QY 209 AAAATGGAGAAATAGCACATATGTGGTAAACACTTCTTCAACTCATTTGGCTTCCCTGTAC 268  
DB 218 CCCAAGGCGCGGACGCAAGTGGATATCGGGTTCGTGCGGACGCGCGCTTCCCGATAC 277  
  
QY 269 TGGTCTCTTTCGGCTTCTTTTCTCGAATCAGCGACCCAAATCAACAGATACAAATTCA 328  
DB 278 TGTCTTCCACCTCTTCTTCTTCTTCCCTCGAAGTACACGCTCCTC-----CT 322  
  
QY 329 GTCACTCCCTTCTTCCACACCCCTTGATCGGTTTACTTGTGCACTGGACTGTAGTGT 388  
DB 323 GCACCAACACTCCCATGCCAAGCTCGCGGTGATATACATCGTCTGGGCTCATCATCG 382  
  
QY 389 CCGCTTATGCTTATTTGTCTCGAGTGGGTTGCTCTACTTACCAAGTCTTACTTTCTCCC 448  
DB 383 CCGCGGACGACATGATGTACATCGTGGCTCAAGTACCTCCCGCTTCGACCTACTCGC 442  
  
QY 449 TCATCTTGGCTCACAGTTGGCTTCACTGCTTTTCTCATATTTCTTAACTCGCAA 508  
DB 443 TCATCTGTGCCAGCCAGCTCGCGTTCAATGCTCTTCTCATACGTGCTCAACTCCAGA 502  
  
QY 509 AGTTCACTCCTTTGATAGTCAGTTCTTTGCTTCTCCTCACTGTATCCTCTCTCTTTG 568

DB 503 AGGTCACTCTCTGATATTCAACTCCGTCTGCTCACCATGTCCTCGCTCGCTCATCG 562  
QY 569 TGGTCAACACTGATTCAGAAAACTCAACTAATGATCTAGAGTACAGTATGTGATCGGGT 628  
DB 563 GAGTCAGCAAGGAGTCTCTCAGGGGGTCAACCGCGTCTCGGAGGAGAGTATCTGCTCGGTT 622  
QY 629 TCATATGTACATCGGTGCTTCCGCTGGGATTGGACCTGTTACTATCTCTGATACAAATGC 688  
DB 623 TCGTGTGACGCTGGGGCGTCTGTCACCTACTCGCTGATCCTCGCGCTGATGACGTCA 682  
QY 689 TCTTCAGGAAAGTTTTTCAAGCAATACATCTCAGCAGTCAAGGACTTGGCCATTATACC 748  
DB 683 CCTTCGAGACCATCATCAAGAGCACACCTTCTCAGCCGTCTCTCAACATGCAGATCTACA 742  
QY 749 AGTCTCTAGTTGGAGTTGTGTCTCTCATAGACCTTTTGTGCAAGTGGAGAGTGGGAAA 808  
DB 743 CGGCGCTCGTGGGACCGGCGCGTCTGGTGGTTCGGGCTGTTTCGGAGCGGAGTGGAGGT 802  
QY 809 CTTTGCCCACTGAGATGAGAAACTACAAACTCGGGAAGTGTATATGTTTGACTTTTAG 868  
DB 803 CGCTGAGGGGAGATGAAACGGTTTCAAGTTCGGGGCAGTCTCTCTACTGATGACGTGC 862  
QY 869 CCTCGGACGCTATTTCTCGCAAGTCTACACTCTCTGCTCTTGTGGGATTTGATCTTCGAGT 928  
DB 863 TGTGGGCGCGGTGTCGTGGCAGTGGCCAAACATCGGGGTGCTCGGCTCATCTTCGAGG 922  
QY 929 CATCTCTGTGTTCTTCCAAATTCATACAGCTGTGGGATTCCTTATAGTTCAGTTCGGG 988  
DB 923 TGTGGGCGCTTCTTCCAACTGATACAGCAGGTGTCACTGCGGTCATCCCGTCTCTTCG 982  
QY 989 CAGTGATAGTTTCCCATGATAGATGGACG 1018  
DB 983 CGGTGGTGGTGTTCACACGACAGGATGAACG 1012

Search completed: November 1, 2004, 07:20:01  
Job time : 632.985 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 1, 2004, 05:45:29 ; Search time 4079.18 Seconds  
(without alignments)  
10228.408 Million cell updates/sec

Title: US-09-913-767-3

Perfect score: 1145

Sequence: 1 ctgtttcatgtgttgatggt.....caagtttaggcaagatcca 1145

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : EST.\*

1: gb\_est1.\*

2: gb\_est2.\*

3: gb\_hic.\*

4: gb\_est3.\*

5: gb\_est4.\*

6: gb\_est5.\*

7: gb\_est6.\*

8: gb\_gse1.\*

9: gb\_gse2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1051.4	91.8	1323	CNS0A36P	BX826811 Arabidops
2	316.8	27.7	1431	CNS0A2KR	BX829081 Arabidops
3	315.6	27.6	1340	CNS0A25J	BX827536 Arabidops
4	312	27.2	1375	CNS0A2S8	BX827187 Arabidops
5	307.8	26.9	1312	CNS0A2U8	BX827470 Arabidops
6	287.2	25.1	1359	CNS0A2R3	BX827477 Arabidops
7	252.4	22.0	1374	CNS0A8BV	BX814955 Arabidops
8	246.6	21.5	782	CO105310	GR_Eb003
9	239	20.9	1519	CNS0A3B5	BX826795 Arabidops
10	231.4	20.2	922	BZ967767	PUDGU56TD
11	229	20.0	770	CG344371	OG4AG157C
12	225	19.7	768	AJ795195	AJ795195
13	220.8	19.3	739	CO105307	GR_Eb003
14	219.4	19.2	837	CB292743	UCRGS01.0
15	218.2	19.1	509	A1729914	BNLGH158
16	217.2	19.0	666	BQ157452	NF105A12I
17	217	19.0	1299	CNS0A41J	BX826394 Arabidops
18	215.6	18.8	1289	CNS0A2H1	BX827065 Arabidops
19	213.8	18.7	633	A1777352	EST363760
20	204.6	17.9	905	BX927581	BX927581
21	203.8	17.8	627	BQ155273	NF078F12I
22	200.2	17.5	630	BG525645	52-8 Stev
23	200.2	17.5	762	BQ121182	EST606758
24	195	17.0	735	BM408950	EST583277

25	194	16.9	806	4	BI310188	BI310188 EST531193
26	190.2	16.6	762	4	BM413141	BM413141 EST587468
27	183.6	16.0	568	4	BM176948	BM176948 saj74h01.
28	183	16.0	701	4	BI204933	BI204933 EST522973
29	182.6	15.9	699	8	B77316	B77316 T3213TF TAM
30	182.6	15.9	800	6	CF069001	CF069001 EST669722
31	180.8	15.8	541	4	BI973110	BI973110 saf40a05.
32	179.8	15.7	676	6	CA920055	CA920055 EST637773
33	179.2	15.7	742	4	BI205696	BI205696 EST523736
34	179.2	15.7	742	4	BI210251	BI210251 EST528291
35	178.2	15.6	882	7	CNI25340	CNI25340 RHOH1.10
36	177.6	15.5	591	4	BI205616	BI205616 EST523656
37	174.8	15.3	593	4	BI432746	BI432746 EST535507
38	174.6	15.2	589	7	CK096760	CK096760 WUND1.39
39	174.6	15.2	765	7	CNI46235	CNI46235 WOUN01.39
40	174.4	15.2	753	4	BM409055	BM409055 EST583382
41	173.6	15.2	691	8	BZ087060	BZ087060 lkh10d1d.
42	173	15.1	757	5	BQ515368	BQ515368 EST622783
43	173	15.1	761	7	CNI25253	CNI25253 RHOH1.10
44	172.2	15.0	562	7	CK759270	CK759270 pam01-13m
45	171.8	15.0	480	4	BG652290	BG652290 sad65a10.

#### ALIGNMENTS

RESULT 1	CNS0A36P	CNS0A36P	1323 bp	mRNA	linear	HTC 06-FEB-2004
LOCUS	Arabidopsis thaliana Full-length cDNA Complete sequence from clone GSU7FB642F05 of Flowers and buds of strain col-0 of Arabidopsis thaliana (thale cress).					
ACCESSION	BX826811					
VERSION	BX826811.1	GI:42460388				
KEYWORDS	HTC; GSLT cDNA.					
SOURCE	Arabidopsis thaliana (thale cress)					
ORGANISM	Arabidopsis thaliana					
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.					
AUTHORS	Castelli, V., Aury, J.M., Jaillon, O., Wincker, P., Clepet, C., Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V., Temple, G., Caboche, M., Weissenbach, J., and Salanoubat, M.					
TITLE	Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences: A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation					
JOURNAL	Unpublished					
REFERENCE	2 (bases 1 to 1323)					
AUTHORS	Genoscope.					
TITLE	Direct Submission					
JOURNAL	Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)					
COMMENT	The sequences are based on single pass reads. Life technologies (a division of Invitrogen) members carried out full-length libraries construction : Temple G. Genoscope members carried out sequencing and annotation : Castelli V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M. URGV INRA : Clepet C., Caboche M. Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich Information center for Protein Sequences). 5 prime and 3 prime are assembled with Phrap. http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full-length http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.					

FEATURES

Location/Qualifiers  
1..1323  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/strain="Col-0"  
/db\_xref="taxon:3702"

Db	955	ACAAAGTTGTGCTGTTCTTCTGCTGTATCTTCTTCCGGTAAAGATAAGTGAATCAAGTT	1014
Qy	984	GGTTGCAATGTTTTTGGCCATCTGGGGAATTTGTTCTTATGTTATCAGCATTTATGTCAA	1043
Db	1015	GGTTGCAATGTTTTTGGCCATCTGGGGAATTTGTTCTTATGTTATCAGCATTTATGTCAA	1074
Qy	1044	TCATAGAAAGCC-AGAAGAAGACC-AAGAGCTTCTCAGTCTTAAAGAAAGAAAGAACAA	1101
Db	1075	TCATAGAAAGCCAAAGAAAGAACCAAGAGCTTCTCAGTCTTAAAGAAAGAAAGAACAA	1134
Qy	1102	AAACAAGTAGATACCAATT-----CATGTCCAAGCTTAGGCCAAAGATCCA	1145
Db	1135	AAACAAGTAAATACCAATATCAAGCGTCCAAGCTTAGGCCAAAGATCCA	1183
RESULT 2			
CNS0A2KR		1431 bp	mRNA linear HTC 06-FEB-2004
LOCUS	Arabidopsis thaliana Full-length cDNA Complete sequence from clone		
DEFINITION	GSLTSL692B03 of Silique of strain col-0 of Arabidopsis thaliana (thale cress).		
ACCESSION	BX829081		
VERSION	BX829081.1		
KEYWORDS	HTC; GSLT cDNA.		
SOURCE	Arabidopsis thaliana (thale cress)		
ORGANISM	Arabidopsis thaliana		
REFERENCE	1 (bases 1 to 1431)		
AUTHORS	Castelli, V., Aury, J.M., Jaillon, O., Wincker, P., Clepet, C., Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V., Temple, G., Caboche, M., Weissenbach, J. and Salanoubat, M.		
TITLE	Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences: A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 1431)		
AUTHORS	Genoscope.		
TITLE	Direct Submission		
JOURNAL	Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)		
COMMENT	The sequences are based on single pass reads. Life Technologies (a division of Invitrogen) members carried out full-length libraries construction : Temple G. Genoscope members carried out sequencing and annotation : Castelli V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M. URGV INRA : Clepet C., Caboche M. Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich information center for Protein Sequences). 5 prime and 3 prime are assembled with Phrap. http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full length http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.		
FEATURES			
source	1. .1431		
	/organism="Arabidopsis thaliana"		
	/mol_type="mRNA"		
	/strain="Col-0"		
	/db_xref="taxon:3702"		
	/clone="GSLTSL692B03"		
	/tissue_type="Silique"		
	/plasmid="pCMVSPORT_6"		
gene	1. .1431		
	/gene="At4g18205"		
ORIGIN			
Query Match	27.7%;	Score 316.8; DB 3;	Length 1431;
Best Local Similarity	57.7%;	Pred. No. 1.4e-73;	
Matches 592; Conservative	0; Mismatches 422;	Indels 12; Gaps 1;	

Qy	28	GAAGGAAATTTTCAACAGAGAGAGAGTCAAGTACTCTTGGAGGTTAAGAGTGTCT	87
Db	197	GAATCTGAATCAITTTTCAGTACCTCAACGGAAGAACTGTAAGAGGTGGCTCCGTGTCTCC	256
Qy	88	CTCTATGTCACTCTCCTCTTAGCTGGAGAGACAATAGCACTCTTTAGTGTAGACTTTTAC	147
Db	257	ATATACGCAATCTTTGTCAATCTTCTGCCAAACACCTTGCTCAGTTCTGGGTGAGACTGTAC	316
Qy	148	TACGAAAAAGCGGTAAAGCACATGGCTCGAAACCTTTGGTTTCAGCTTGTAGGGTTTCCT	207
Db	317	TATGAAATGAGGGAAGACATATGTGGTAAACACTTCTTCAACTCATTTGGCTTCCT	376
Qy	208	TTAACCCTCTCTGTCTATTAATTAATAAGCCTGAGCGTCCAAGACTTAAACCACTTACC	267
Db	377	GTACTGATTTCTTCCGCTTCTTTCTCGAATCAGGCCAACCCAAATCAACAGATACAAT	436
Qy	268	AAAAAACTACTTCTTCTCTTGACACTATCTTTAGTGTATATTTGGACTTTGGCTTGGTT	327
Db	437	TTCAGTCAGTCCCTTCTTCCACCCTTCGATCGGTTTACTTGTGCACGTGSACTGCTA	496
Qy	328	GTTGCTGGACATGTATTTTGTACTCATTTTGGGTACTTTTACCTTCTCTGTCTCAACTTTC	387
Db	497	GTGTCTGCTTATGTCTTATTTGTCTGAGTTGGTGTCTTTACTTACCACTCTCTACTTTC	556
Qy	388	TCCTTGATCTCTGGCTCGCAATTTGGCTTTTAAACCGCTTCTCTTACTTCTCTAACTCA	447
Db	557	TCCTCATCTTGGCTCACAGTTGGCTTCTACTGCTTTTTCTCATATTTCTTTAACTCG	616
Qy	448	CAAAAAATCACACATTTATTACTCAATTCACCTTGTCTTTAAACCAATCTTCTTACACTT	507
Db	617	CAAAAGTTCACTCTCTTTGATAGTCAATTTCTTTGTCTCTCTTACGGTTTCTCTGCGCTC	676
Qy	508	CTTGTTATCCAACATGAACCAAGATCTCCCTCTTCTACTTCAAAGTCGGACGCCAAGTCC	567
Db	677	CTCGTGTCAACACTGATTCAGAA-----AACACAACAATGTATCTAGAGTA	724
Qy	568	AAGTATGTGATTTGGATACATCTGGCGGTTCGGTAGCTCAGCTGTGTATTTCTCTGGTGT	627
Db	725	CAATATGTGATTTGGTTTCATCTGTACCATTTGCTGCTTCCGCTGGGATTTGGTACTA	784
Qy	628	TCCTTTACAGATTAAGCGTTTCGAAAAGATTTCAAAGAAATACACATTCAAAGGCTATT	687
Db	785	TCTCTGATACAACCTGCTCTTCAGGAAAGTTTTCAGAAAGCATACATCTCTACGACGCTG	844
Qy	688	GACATGGCCATATCCGTCTATGGTAGTACTTTGTGTAGTTGTGGTAGGACTTTTGGGA	747
Db	845	GACTTGGCCAAATACCACTCTCTAGTTGCTACTTTGTGTGTACTCATAGGACTGTGTTGCA	904
Qy	748	AGTGGTGGGTGAAAAAGCTCAGTATACAGAAATGGAAGTTTCAACTAGGAAAAAGTCA	807
Db	905	AGTGGAGTGGAGATTTCTGCCAAGTGAGATGAGAACTACATATCTGGGGAAGTGTCA	964
Qy	808	TACATTTTGAATAACATCGGTTCAACGATATCATGGCAAGCTTGTTCATTGGAAGTGT	867
Db	965	TATATCTGACTTTTGGCTCAGCAGCTATTTTCTGGCAAGTATACACTGTTGGTTGTGTG	1024
Qy	868	GGTTTGATATCGAAGTTTCACTCGCTTTTTCGAAATGTCAATAGCACTCTTTGTTTACCA	927
Db	1025	GGATTAACTCTTCGAGTCTTCTTCTGTGTCTTCCAAATTCCTAACTGCTGTGGGACTG	1084
Qy	928	GTTCTGCTGTCTTTGTGTGTGTCTTCTTCGCTGATGAGATGAGTGGAAATCAAGTGTGT	987
Db	1085	ATAGTTCAGTTGTAGCAGTGAATAGTTTTCATGATAGATGGAATGCAATCAAGATTTTC	1144
Qy	988	GCAATGTTTTTGGCCATCTGGGGAATTTGTTTCTTATGTGTTATCAGCAATATGTCAATGAT	1047
Db	1145	TCCATCATTTTAGCTATCTGGGGCTTCTCTTTCAATTTGTCTATCAGAACTACCTCGACGA	1204
Qy	1048	AGAAAG	1053
Db	1205	AAGAAAG	1210

RESULT 3	CNS0A29J	1340 bp	mrna	linear	HTC 06-FEB-2004
LOCUS	CNS0A29J				
DEFINITION	Arabidopsis thaliana Full-length cDNA Complete sequence from clone ASLTLS71ZA05 of Adult vegetative tissue of strain col-0 of Arabidopsis thaliana (thale cress).				
ACCESSION	BX827536				
VERSION	BX827536.1	GI:42459416			
KEYWORDS	HTC; GSLT cDNA.				
SOURCE	Arabidopsis thaliana (thale cress)				
ORGANISM	Arabidopsis thaliana				
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.				
AUTHORS	Castelli,V., Aury,J.M., Jaillon,O., Wincker,P., Clepet,C., Menard,M., Cruaud,C., Queier,F., Scarpelli,C., Schachter,V., Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.				
TITLE	A Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences: A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 1340)				
AUTHORS	Genoscope.				
TITLE	Direct Submission				
JOURNAL	Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)				
COMMENT	The sequences are based on single pass reads. Life Technologies (a division of Invitrogen) members carried out full-length libraries construction : Temple G. Genoscope members carried out sequencing and annotation : Castelli V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M. URV INRA : Clepet C., Caboche M. Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich Information center for Protein Sequences). 5 prime and 3 prime are assembled with Phrap. http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full-length				
FEATURES	source				
ORIGIN	http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis. Location/Qualifiers				
	1..1340				
	/organism="Arabidopsis thaliana"				
	/mol_type="mRNA"				
	/strain="Col-0"				
	/db_xref="taxon:3702"				
	/clone="GSLTLS71ZA05"				
	/tissue_type="Adult vegetative tissue"				
	/plasmid="pCMVSPORT_6"				
gene	1..1340				
	/gene="At4g18197"				
ORIGIN					
Query Match	27.6%	Score 315.6;	DB 3;	Length 1340;	
Best Local Similarity	57.4%	Pred. No. 3e-73;			
Matches	595;	Conservative	0;	Mismatches 429;	Indels 12; Gaps 1;
Qy	18	GGTAGAACCTGAGGGGAAATTTTCAACAGAGAGAGAAAGTCAACAGTACTCTCTGGAGGTT	77		
Db	103	GGTGGTAACCTGAATCATCATCATCAGCTGTGCGCTCAAAACCGAGAACTATATAAAGGTGGCT	162		
Qy	78	AAGAGTGTCTCTATGTCTACTCTCTCTAGCTGGAGAGACAATAGCCACTCTCTTAGG	137		
Db	163	TCGTGTCTCCATATACGTAAATCTTTGTGCTCTTTTGCACGCCACTAGTACAAATCTGGG	222		
Qy	138	TAGACTTTTACTAGCAAAAAGCGGTAAAAACATGGCTCGAAACCTTGGTTCAGCTTGT	197		
Db	223	TAGATTGTACTATGAAATGAGAAATAGCACATATGTGGTAACACTTCTTCACTCAT	282		
Qy	198	AGGGTTTCCTTTAACCCCTTCCTTGCTATTAATTACTTAAAGCCTGAGCCGTCGAAGACTAA	257		
Db	283	TGGCTTCCCTGTACTGGTCTCTGTTCCGCTCTCTTTCTCGAATCAGGCAACCCAAATCAAC	342		





Db 817 TCATAGGACTGTTGCAAGTGGAGTGGAGAACTCTGCCAAGTGAGATGAGAACTACA 876  
Qy 791 AACTAGGAAAGCTCATACATTTTGTGATAACATCGGTTCAACGATATCATGCGAAGCTT 850  
Db 877 AACTGGGAAAGTGTATATATCTTGACTTTGGGCTCAGCAGCTATTTCTTGGCAAGTAT 936  
Qy 851 GTTTGATTGGAAGTGTGGTTGATTATTCGAAAGTTTCATCGCTTTTTCCTCAATGCTATA 910  
Db 937 ACACGTGTTGGTGTGGGATTAACTTCGAGTCTTCTTCTGTGTTCTCCAAATTCCTATA 996  
Qy 911 GCACCTTTTGTACCAGTGTGCTGTCTGTGTGTGTCTTCTCGTGATCAGATGA 970  
Db 997 CTGCTGGGACTGCCTATAGTTCAGTTGTAGCAGTATGATGTTTTCATGATGAAGTGG 1056  
Qy 971 GTGGAATCAAGTGTGGTTGCAATGTTTTTGGCCATCTGGGGATTTGTTTCTTATGTTATC 1030  
Db 1057 ATGCATCAAGATTTTCTCCATCATTTTAGCTAICTGGGGCTTCTTCAITTTGCTATC 1116  
Qy 1031 AGCAATTATGTCATGATGAAG 1053  
Db 1117 AGCACTACCTCGACGAAAGAAG 1139

RESULT 6  
CNS0A2R3 1359 bp mRNA linear HTC 06-FEB-2004  
LOCUS Arabidopsis thaliana Full-length cDNA Complete sequence from clone  
DEFINITION GSLT55ZE07 of Adult vegetative tissue of strain col-0 of  
Arabidopsis thaliana (thale cress).

VERSION BX827477.1 GI:42459927  
KEYWORDS HTC; GSLT cDNA.  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
1 (bases 1 to 1359)  
Castelli,V., Aury,J.M., Jallou,O., Wincker,P., Clepet,C.,  
Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V.,  
Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.  
Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences:  
A Combined Approach to Evaluate and Improve Arabidopsis Genome  
Annotation

JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 1359)  
AUTHORS Genoscope.  
TITLE Direct Submission  
JOURNAL Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :  
BP 191 91006 EVRY cedex - FRANCE (E-mail : [segref@genoscope.cns.fr](mailto:segref@genoscope.cns.fr))  
COMMENT - Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr)  
The sequences are based on single pass reads.  
Life Technologies (a division of Invitrogen) members carried out  
full-length libraries construction : Temple G.  
Genoscope members carried out sequencing and annotation : Castelli  
V., Aury J.M., Jallou O., Wincker P., Menard M., Cruaud C.,  
Schachter V., Weissenbach J., Salanoubat M.  
URGV INRA : Clepet C., Caboche M.  
Annotation is based on the June 2003 version of the Arabidopsis  
genome released by MIPS (Munich Information center for Protein  
Sequences). 5 prime and 3 prime are assembled with Phrap.  
[http://www.genoscope.cns.fr/externe/sequences/Banque\\_Projet\\_EP/Full\\_length](http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EP/Full_length)  
<http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis>.

FEATURES  
source  
1. 1359  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/strain="Col-0"  
/db\_xref="taxon:3702"  
/clone="GSLT55ZE07"  
/tissue\_type="Adult vegetative tissue"

gene /plasmid="pCMVSPORT\_6"  
1. 1359  
/genes="At4g18197"  
ORIGIN  
Query Match 25.1%; Score 287.2; DB 3; Length 1359;  
Best Local Similarity 56.9%; Pred. No. 1.2e-65;  
Matches 591; Conservative 0; Mismatches 433; Indels 14; Gaps 3;  
Qy 18 GGTAGAACCTGAAGGGAAATTTTCAACAGAGAGAGAAAGTCACAAAGTACTCTTTGGAGGTT 77  
Db 77 GGTGGTAACCTGATCATCATCATCAGCTGTGCCTCAAAACGAGAACTATATAAAGGTGGCT 136  
Qy 78 AAGAGTGTCTCTATATGTCACTCTCCTCTTAGCTGGAGAGACAATAGCCACTCTCTTAGG 137  
Db 137 TCGTGTCTCCATATAGATGAATCTTTGTCTCTTTTGGCAGCCACTAGTCAAAATCTCGG 196  
Qy 138 TAGACTTTTACTAGCAAAAAGCGGTAAAGACATGGCTCGAAACCTTGGTTCAGCTTGT 197  
Db 197 TAGATTGTACTATGAAAATGAGAGAAATAGACATATGTGTGAACACTTCTTCAACTCAT 256  
Qy 198 AGGTTTTCCTTTAAACCTTCTTGTCTATTAATTAATTAAGCCCTGAGCGCTCCAAGATA 257  
Db 257 TGGCTTCCCTGTACTGCTTCTTCTCGAATCAGGCAACCCAAATCAAC 316  
Qy 258 AACCATTTACAAAATACTACTTCTTCTTCTTGGACACTATCTTATAGTGATATGGACT 317  
Db 317 AGATACAAATTTGAGTCAGTCCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 376  
Qy 318 TGGCTTCTTGTGTGGACATTTGATTTTGTACTCATTTTGGGCTACTTTTACCTTCTTCTTCT 377  
Db 377 TGGACTGTAGTGTCCGCTTATGCTTATTTGTCTGACGATAGGTTGCTCTACTTACCACT 436  
Qy 378 CTCAACTTCTTCTTGTATCTCTGCGTGCATTTTGGCTTTTAAAGCCGCTTCTTCTTCTTCTT 437  
Db 437 CTCTACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 496  
Qy 438 CCTAACTCTCAAAAATACTACACCATTTTATATCTTCAATTTCTTCTTCTTCTTCTTCTTCTT 497  
Db 497 CCTTAACTCGCAAAAGTTTCACTCTTGTAGTAGTGTCTTGTCTTCTTCTTCTTCTTCTTCTT 556  
Qy 498 TTCTACACTTCTTGTATTCAACATGAACCAAGATCTCCCTCTTCTTCTTCTTCTTCTTCTTCTT 557  
Db 557 CTCTGCTTCTTGTGTGCTCAACTGATTCAGAA-----AACTCAACTAATGT 604  
Qy 558 AGCCAAGTCCAAGTATGTGATGTGATATCATCTGCGCGGTGCGGTAGCTCAGCTGTGTTATTC 617  
Db 605 ATCTAGAGTACAGTATGTGATCGGGTTTCATATGTACCATCGGTCTTCCGCTGGGATTGG 664  
Qy 618 TCTGGTCTTCTTCTTAAACAGATTACGGTTTCGAAAGATTTCTAAAGAAATACACATTCAA 677  
Db 665 ACTGTACTATCTCTGAGACAAATGATCTTTCAGAAAGTTTTTTCAGAAAGATACATCCTC 724  
Qy 678 GGCTATTTTATAGACATGGCCACATATCGCTCTATGTTAGTACTTGTGTAGTGTGGTAGG 737  
Db 725 AGCAGTAACGGACTTGGCCATTTTACCACTCTCTAGTTGCGGTTGTGTAGTTCTCATAGG 784  
Qy 738 ACTTTTGGAAAGTGTGGTGGTGGGAAAGAGTGTGATACAGAAATGGAAGAGTTTCAACTAGG 797  
Db 785 ACTTTTGGAAAGTGTGAGTGGGAAACTTTTGGCCACGTGTGATGAGAAACTACAAACTCGG 844  
Qy 798 GAAAGCTCATACATTTTGTATTAACATCGGTTCACGATATCATGCAAGCTTGTGTTGAT 857  
Db 845 GAAAGTGTATATGTTTGTGATTAAGCTCGGCAACTATTTTTCATGGCAAGTCTACACTCT 904  
Qy 858 TGGAAAGTGTGG-TTTTGATTATCGAAGT-TTCACTCGCTTTTTTCCAAATGTCAAGCACT 915  
Db 905 TGGTCTTGTGGGATATGATCTTTCGAGTTCATCTCTGTGTTTACCCTTCCATCCCGCT 964  
Qy 916 CTTTGTTTACAGTTGTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 975  
Db 965 GTGGGCATGCTATCGTTCCTTCCGTTGGCGCGGTATCGTTTCTTCTTCTTCTTCTTCTTCTTCT 1024







Db 254 AACCAGTTGCTACATTTTGGGCGAGAGTATATATGACAAACGGAGGAAACAGTAATGCG 313  
Qy 176 TC-GAAACCTTGCTGAGCTGTAGGTTTCCCTTTAAACCTTCCCTTGCTATTATTACTTA 234  
Db 314 TTAGCAACGCTAGTCTCAACTGTTGGCTTCCCTGTTCTACTTCCATATTATATCTGTCA 373  
Qy 235 AAGCCTGAGCGGTCAAGACTAAACATTAACCAAAAAAACAATCTTCTTCTTCTTGACA 294  
Db 374 TTTAAAAACATGCAACAACTGATAGAGATGGAAAAAGAACCTCACCTAG-----GAAC 427  
Qy 295 CTATCTTTAGTGATATTGGACTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 354  
Db 428 CGTGATTTGGTTTACGTAGTCTGGACTTCTTGAGGAGCAGATTGCTATTGTACTCC 487  
Qy 355 TTTGGGCTACTTACCTTCTGCTCAACTTCTCTTTGATCTCTGCGTCGCAATTTGGCT 414  
Db 488 ATTGGACTTCTTTACTTACCCTTTCTACCTATTCCCTGATCTGTGCACTCAGITAGCC 547  
Qy 415 TTTAAGCGCGTCTTCTTACTTCTTAACTCAAACTCAAAAAAATCACACCATTTATCTCAAT 474  
Db 548 TTAATGCTTTTCTTCTTATTTCTTAACTCAAAAAAATTTACCCCTATCATTTTAAAT 607  
Qy 475 TCACCTTGTCTCTTAACCATATCTCTACACTTCTTGTATCCAAACATGAACAGATCT 534  
Db 608 TCTCTTTCTCTAACTATATTTTCCACCTACTTGGCATTCATATAGGAGAGAGAT 667  
Qy 535 CCTCTTCTTACTTCAAGTCCGAGCAAGTCAAGTATGTGATTTGATGATCATCTCGCGG 594  
Db 668 TCCACAAAAGTTACAAAAGGAG-----ATTTTTAAAGTTTTCATATGCACC 715  
Qy 595 GTGCGTAGCTAGCTGGTTATCTCTGGTCTTCTTCTTAAACAGA-TTACGGGTTCGAAA 653  
Db 716 GTTGTGCTTTTGTCTGTTATGTTTATTTTATCCCTACAACAGCGAGCCTTTTAA 775  
Qy 654 GATTTAAAGAATACACATTCAGGCTATTTTAGACATGGCCACATATCCGCTCTATGT 713  
Db 776 AGTCTTAAAGAGCAAAATTTCTCAAGTTATGGATATGATATCTACATGAATCTGAT 835  
Qy 714 AGCTACTTGTGTAGTTGTGTAGGACTTTTGGAGTGGGTGGGTAAGGAGTCAAGTAC 773  
Db 836 GCGAGTTGTGTTCCGCTGTGGGCTTTTGTCTACAATGAATGGAAAAATTTGAACAA 895  
Qy 774 AGAATGGAAGATTTCAACTAGGAAAAAGCTCATACATTTTGTATAAACATCGTTTCAAC 833  
Db 896 TGAAGAGGAAAACTAAAAAAGGAAAGTAACCTACAATATGAACCTAATGTGAACAGC 955  
Qy 834 GATATCATGCAAGCTTGT--TTGATTGGAAGTTGTTGTTGATATCGAAGTTTCATCG 891  
Db 956 TGTACTCTGCGCGTATCTCCATCGGTGGAAACAAGAACTGAATCTTCGAGCTCTCCTCT 1015  
Qy 892 CTTTTTTTCCAATGTCAT-AAGCACTCTTTGTTTACCAGTTGTGCGCTGTTCTTGCTGTTGT 950  
Db 1016 CTATCTTAAATGCAATAAAGGTTTGGAACTCCAGTGGTTCCTATCTTGCTGTAT 1075  
Qy 951 CTTCTTCGTGATGAGATGAGTGGAAATCAAGTTGGTGTGCAATGTTTTTGGCCATCTGGGG 1010  
Db 1076 CATTTTCCATGACAAAATAAATGGGTAAAGGTGATTTCTATGATCTAGCTATTGGGG 1135  
Qy 1011 ATTTGTTTCTTATGTTATCAGCATATGTCATGATGATGAAAGCCAGAGAGACAAGA 1070  
Db 1136 TTTAACTTCTTATGCTACCAAAAATATCTTGTATGACAAAAAATTTGAAGAAAAATCATGA 1195  
Qy 1071 GCTTCTCAGTCTAAAGAGAGAGAAAGAAACAAAAACAAGTAGATACCAAT 1119  
Db 1196 AATAACAAACAGATATCCCTGACCCACAGAGAGAGAGTCAACT 1244

RESULT 10  
B2967767  
LOCUS  
DEFINITION  
PUDG56TD ZM\_0.6\_1.0\_KB Zea mays genomic clone ZMMBTa184115,  
genomic survey sequence.  
ACCESSION  
B2967767

VERSION B2967767.1 GI:29185853  
KEYWORDS GSS.  
SOURCE Zea mays  
ORGANISM Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.  
REFERENCE 1 (bases 1 to 922)  
AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,  
Reinick,A., Praser,C.M., Yuan,Y., San Miguel,P., Ma,J. and  
Bennetzen,J.  
Maize Genomics Consortium  
Unpublished (2003)  
CONTACT: Cathy Whitelaw  
TIGR  
7912 Medical Center Drive, Rockville, MD 20850, USA  
Tel: 301-838-5843  
Fax: 301-838-0208  
Email: whitelaw@tigr.org  
Seq primer: 7F  
Class: sheared ends.  
FEATURES  
Location/Qualifiers  
1..922  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/strain="B73"  
/db\_xref="taxon:4577"  
/clone="ZMMBTa184115"  
/clone\_lib="ZM\_0.6\_1.0\_KB"  
/note="Vector: PCR4-TOPO; Site\_1: EcoRI; 0.6-1.0 kb high  
Cor selected genomic DNA library"  
ORIGIN  
Query Match 20.2%; Score 231.4; DB 8; Length 922;  
Best Local Similarity 58.3%; Pred. No. 1.1e-50;  
Matches 431; Conservative 0; Mismatches 296; Indels 12; Gaps 1;  
Qy 317 TTGGCTTCTTGTGCTGGACATTTGTTGTTACTCAATTTGGGCTACTTTTACCTTCTCTG 376  
Db 3 TGGGCTCATCTGCTGCCGATGACTTGATGTTATCTATGGCCTTTTGTACCTTCTG 62  
Qy 377 TCTCACTTCTCTTTGATCTCTGGTTCGAAATTTGGCTTTTAAACGCCCTTCTCTTACT 436  
Db 63 TATCAACATATTTCGCTCATTTTGGCTAGTACGTGGCCTTCAATGCTGCTTCTCATATG 122  
Qy 437 TCCTAACTCAAAAAATCACACATTTATCAATTTACTCAATTTGCTTCTTAAACATAT 496  
Db 123 TCCTAAATGCTCAAAAGTTTACCCCATTTCACTTTTCACTCAGTAATTTCTCTTACCTTT 182  
Qy 497 CTTCTACACTTCTTGTATTCCAAACATGAACAGAACTCTCCCTCTTCTACTTCAAAAGTCG 556  
Db 193 CGCTGCGCTTCTGGAGTTGACGAGA-----TTCTCAGGTTACCAATGGTT 230  
Qy 557 CAGCAAGTCCAAAGTATGATTTGATATACATCTGCGCGGTTCGGTAGCTCAGCTGGTTAT 616  
Db 231 TATCGCGTGGGAAGTACATATTGGGTTTTCGCAATTTGACCTAGGAGCCTTCGGCCACAT 290  
Qy 617 CTCTGGTCTTCTTTTAAACATTTACCGTTTCGAAAGATTTCTAAAGAAATACACATCA 676  
Db 291 CACTAATTTCTCTCTAATGCAAGTCGCATTTTCGAGAAGGTTATTAAAGAAAGAACTTTCT 350  
Qy 677 AGGCTATTTTACAGATGGCCACATATCCGTCTATGGTAGTACTTGTGTAGTTGGTAG 736  
Db 351 CAGTCGTGTTGAATATGAGATATATACAGCACTAGTGGCAACAGTAGTCTTCTTATCG 410  
Qy 737 GACTTTTGGAGTGGTGGGAAAGTGAAGTGAAGTACAGAAATGGAAGATTTTCAACTAG 796  
Db 411 GTTTATTTCGAAGCGCGAGTGAAGACTTTAGAGGGAGAGATGATGCCTTTCAGCTCAG 470  
Qy 797 GAAAAAGCTCATACATTTTGTATAAACATCGGTTTCAACGATATCATGGCAAGCTTGTGA 856  
Db 471 AGAGGGTGTCTGTATGTATGACATTTCTATGGACTGCTGTATCTTGGCAGATAGCTCCG 530



Best Local Similarity		57.6%	Pred. No. 5.2e-49;	Matches 452; Conservative 0; Mismatches 315; Indels 18; Gaps 2;	
Qy	160	GGTAAAGCACAATGGCTCGAAACCTTGGTTTCAGCTTGTAGGGTTTCCCTTAAACCTTCCT	219		
Db	2	GGGAACAGCAATGGCTAGCCACATAGTACAAGTTATAGACTTTCNAATCTCGTCCCA	61		
Qy	220	TGCTATTATTACTTAAAGCCTGAGCGGTCGAAGACTAAACCAATTACCAAAAAAACTACT	279		
Db	62	TTCAAATGGTTTCAACAACACAGAAACGAGGGCTAATGAACCAAAACACAACAAGA----	116		
Qy	280	TCITCCCTCTTGACACTATCTTTAGTGTATATTGGACTTGGCTTGCCTTGTGTGCTGGACAT	339		
Db	117	-AATCCCCCACTGGTCTTGTATTATATATGTAGTCTCGGCATTTTGTGCGCAGGAAT	175		
Qy	340	TGTAATTTTGTACTCAATTTGGGCTACTTTTACCTTCTCTGTCTCAAATTTCTCTTTGATCTCT	399		
Db	176	TCCATGTATACACAATTTGGACTCCATACCTTCGGGTGACTACGTACACTTTGATCTGCG	235		
Qy	400	CGGTGCGAATTTGGCTTTTAAACGCGCTCTTCTTACTTCTTAAACTCACAAAAAATCACAA	459		
Db	236	GCTAGTCGGTTAGGATTCACGCGCTTTTCTCGTTTTTCCCTCAACAAACAAAAGTTCACT	295		
Qy	460	CCATTTATACTCAATTCACCTTGTCTCTTAACCATATCTTCTACACTTCTTGTATTCCAA	519		
Db	296	CCITTACATTTAAACTCTGTAGTACTCCTTACGANTAGTTCAAATATTACTAGTGTTCCAA	355		
Qy	520	CATGAACAGAAATCTCCCTCTTCTACTTTCAAAGTCCGAGCCAAAGTCCAAGTATGTGATT	579		
Db	356	ACTGATCTGGAGACTCCCAAAAGCTCGAAAA-----AAAGTATGTCAATT	403		
Qy	580	GGATACATCTCGCGGCGGTAGTCTAGCTGGTTATTCTCTGGTGTCTTTCTTTAACAGAT	639		
Db	404	GGCTTTTATCCACTGTGCGCGCCTCTGCTGGCTACGCGTTAGTCTTTTCAATCACACAG	463		
Qy	640	TACGGGTTCGAAAGATTTCTAAGAAATACACATTCAGGCTATTTTAGACATGGCCACA	699		
Db	464	CTCGCCTTCGTAAGGTCATCAAGAAATGAACCCATTCGCGCGATTCCTGGAAATTCACAATC	523		
Qy	700	TATCCGCTCATGGTAGTACTTGTGTAGTTGTGTGTAGGACTTTTGTGAAGTGGTGGGTGG	759		
Db	524	TATCAGAAATTTGGTGGCGACTTTCGTGATCCTAATAGCCCTTTTTCGCAAGTAGGAGTGG	583		
Qy	760	AAAAAGCTGAGTACAGAAATGGAAGAGTTTCACTAGGGAAGAGCTCATACATTTTGATA	819		
Db	584	AAGGATTTGAACAAAGAAATCGAGGGGTTTAAAGTTGGGGAAGTTTCTTACATCATGAAC	643		
Qy	820	AACATCGGTTCAACGATATCATGCGCAGCTTGTTTGATTGGAAGTGTGGTTTGAATTATC	879		
Db	644	TTGTTTTGGACAGCAGTGTCTGTGGCAGACGTTAATGTCTGGTGTATTTGGATTGATCTTC	703		
Qy	880	GAAGTTTTCATCGCTTTTTTCCAAATGTATAGCACTCTTGTGTTTACAGTGTGTCCTGTT	939		
Db	704	AAGATTTCCGGCTTGTGTTTTCGAAATGTGATAGTATTTTAGGCCTTCGTTGGCTCCGGTT	763		
Qy	940	CTTGC	944		
Db	764	CTTGC	768		

RESULT 13	CO105307	739 bp	mRNA	linear	EST 16-JUN-2004			
LOCUS	GR_Eb0035G13.r	GR_Eb	Gossypium	raimondii	cDNA clone GR_Eb0035G13			
DEFINITION	3', mRNA sequence.							
ACCESSION	CO105307							
VERSION	CO105307.1	GI:48803993						
KEYWORDS	EST.							
SOURCE	Gossypium raimondii							
ORGANISM	Gossypium raimondii							
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eursoids II; Malvales; Malvaceae; Malvoideae; Gossypium.								





**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 1, 2004, 06:20:04 ; Search time 559.891 Seconds  
(without alignments)  
9900.026 Million cell updates/sec

Title: US-09-913-767-6

Perfect score: 1081  
Sequence: 1 caaatccacagttcaagat.....ataagctgattttgaaag 1081

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3413475 seqs, 2563800928 residues

Total number of hits satisfying chosen parameters: 6826950

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.\*  
1: /cgn2\_6/prodata/1/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/prodata/1/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/prodata/1/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/prodata/1/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/prodata/1/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/prodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/prodata/1/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/prodata/1/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/prodata/1/pubpna/US09A\_PUBCOMB.seq.\*  
10: /cgn2\_6/prodata/1/pubpna/US09C\_PUBCOMB.seq.\*  
11: /cgn2\_6/prodata/1/pubpna/US09\_NEW\_PUB.seq.\*  
12: /cgn2\_6/prodata/1/pubpna/US10A\_PUBCOMB.seq.\*  
13: /cgn2\_6/prodata/1/pubpna/US10B\_PUBCOMB.seq.\*  
14: /cgn2\_6/prodata/1/pubpna/US10C\_PUBCOMB.seq.\*  
15: /cgn2\_6/prodata/1/pubpna/US10D\_PUBCOMB.seq.\*  
16: /cgn2\_6/prodata/1/pubpna/US10E\_PUBCOMB.seq.\*  
17: /cgn2\_6/prodata/1/pubpna/US10F\_PUBCOMB.seq.\*  
18: /cgn2\_6/prodata/1/pubpna/US10\_NEW\_PUB.seq.\*  
19: /cgn2\_6/prodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
20: /cgn2\_6/prodata/1/pubpna/US60\_NEW\_PUB.seq.\*  
21: /cgn2\_6/prodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	406.2	37.6	3387	9	US-09-938-842A-2212
2	406.2	37.6	3387	11	US-09-938-842A-2212
3	361.8	33.5	1294	16	US-10-424-599-2524
4	307	28.4	1513	17	US-10-437-963-82061
5	303.8	28.1	2214	16	US-10-424-599-102519
6	262	24.2	2381	17	US-10-437-963-43860
7	234.2	21.7	83698	17	US-10-416-898-9
8	223.4	20.7	997	16	US-10-424-599-75161
9	179.2	16.6	728	17	US-10-767-701-9000
10	149	13.8	522	16	US-10-424-599-114931
11	141.6	13.1	1386	16	US-10-425-114-12485
12	138.6	12.8	592	9	US-09-770-152-44

13	138.4	12.8	1779	16	US-10-425-114-28473
14	131.8	12.2	1233	16	US-10-424-599-30274
15	131	12.1	1146	16	US-10-425-114-10516
16	129.2	12.0	2605	17	US-10-437-963-83900
17	127	11.7	1432	17	US-10-437-963-1897
18	126.4	11.7	332	11	US-09-732-627A-1196
19	125.8	11.6	671	17	US-10-767-701-5986
20	123	11.4	1373	16	US-10-424-599-31364
21	117.6	10.9	792	16	US-10-424-599-9213
22	115.2	10.7	1333	16	US-10-425-114-10574
23	109.6	10.1	1355	16	US-10-424-599-11363
24	108.2	10.0	327	16	US-10-424-599-121152
25	102.6	9.5	650	17	US-10-767-701-4045
26	97.4	9.0	747	17	US-10-767-701-6806
27	82.2	7.6	994	16	US-10-424-599-73501
28	73.8	6.8	308	16	US-10-424-599-121805
29	72.2	6.7	774	9	US-09-770-445-895
30	71.4	6.6	1173	17	US-10-437-963-57539
31	66.2	6.1	386	9	US-09-770-791-85
32	63.2	5.8	716	16	US-10-424-599-61310
33	60.4	5.6	1417	16	US-10-424-599-97043
34	59.6	5.5	2928	17	US-10-437-963-33800
35	57	5.3	789	17	US-10-437-963-72452
36	56.4	5.2	1177	17	US-10-437-963-91425
37	55.4	5.1	382	17	US-10-021-323-592
38	55.4	5.1	383	17	US-10-021-323-2053
39	42.6	3.9	59475	17	US-10-322-696-166
40	42.2	3.9	201	17	US-10-741-601-16181
41	42.2	3.9	81210	17	US-10-741-601-5664
42	42	3.9	357652	17	US-10-322-696-34
43	41.8	3.9	1296	16	US-10-425-114-13687
44	41.4	3.8	684187	17	US-10-367-094-71
45	40.6	3.8	72604	15	US-10-162-497-7

ALIGNMENTS

RESULT 1  
US-09-938-842A-2212  
; Sequence 2212, Application US/09938842A  
; Patent No. US20020160378A1  
; GENERAL INFORMATION:  
; APPLICANT: Harper, Jeff  
; APPLICANT: Kreps, Joel  
; APPLICANT: Wang, Xun  
; APPLICANT: Zhu, Tong  
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
; FILE REFERENCE: SCRI3300-3  
; CURRENT APPLICATION NUMBER: US/09/938,842A  
; CURRENT FILING DATE: 2001-08-24  
; PRIOR APPLICATION NUMBER: US 60/227,866  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: US 60/264,647  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/300,111  
; PRIOR FILING DATE: 2001-06-22  
; NUMBER OF SEQ ID NOS: 5379  
; SEQ ID NO 2212  
; LENGTH: 3387  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-09-938-842A-2212  
Query Match 37.6%; Score 406.2; DB 9; Length 3387;  
Best Local Similarity 64.7%; Pred. No. 2.2e-104;  
Matches 622; Conservative 0; Mismatches 333; Indels 6; Gaps 1;  
QY 88 GGAGTATCTATGCAACGGAGGAAACAGTAATGGCTAGCAACGGTAGTTCACCTGTT 147  
DB 236 GTAGTTTACTATGAAATGGTGGGAAGAGATCATGGATGGGAACACTTGTCCAACTATC 285



Db 946 CTGATCTTTGAGTCATCTTCTGTGTTCTCAATTCATTAACCTCTCTGGGATTCGCTATA 1005  
Qy 862 GTTCCTATCTTGCTGTATATCATTTTCCATGACAAATGAATGGGTAAAGGTGATTTCT 921  
Db 1006 GTTCCAGTTGTAGCAGTGATTTGTTTCCATGATAAATGAACGGTCAAAGATCTTCTCC 1065  
Qy 922 ATGATCCTAGCTATTTGGGGTTTCACTTCTCTATGTTCTACCAAAATATCTTGTATGACAA 981  
Db 1066 ATCATTTTAGCTATCTGGGGATTCATTTCTATTTGTCTATCAGCATACCTCGACGAAAAG 1125  
Qy 982 AACTTGAAGAAAATCATGAAATCACAACAAAGAAATCCCTGACCCACAGAGCAGAA 1041  
Db 1126 AAGTTGAAGACTAGCCACACAGTCTCTGTAGGAGATCCTCATCTACTACTGCTGAGAA 1185  
Qy 1042 G 1042  
Db 1186 G 1186

## RESULT 3

US-10-424-599-2524  
; Sequence 2524, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 2524  
; LENGTH: 1294  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_102285C.1  
US-10-424-599-2524

Query Match 33.5%; Score 361.8; DB 16; Length 1294;  
Best Local Similarity 63.7%; Pred. No. 5.6e-92;  
Matches 596; Conservative 0; Mismatches 312; Indels 27; Gaps 2;  
Qy 87 AGGAGTATCTATGACACCGGAGAACAGTAAATGCGTAGCAACGGTAGTTCAACTTGT 146  
Db 199 AGATTGTACTATGAAGAGGTGGAAAAAGCAAGTGGAGAACACTTGTTCAACTTGC 258  
Qy 147 TGGCTTTCCTGTCTACTTCCATATTATATCTTGTCAATTTAAACACATGCAACAACTGA 206  
Db 259 TGGTTTCCCTATTTCAGCTTCTTTCACCTTATTTTCAGCATCCAAAATCTCACCACAA 318  
Qy 207 TAG-----AGATGAAAAGAACCTTACCTAGGAACCGGTGATTTGGTTTAGCTAGTGT 260  
Db 319 TAGTAGCATTCATCCAAAACAATCATCTGCTTCAATACTAGCATTTATCTATGCTCAAT 378  
Qy 261 TGGACTTCTTGTAGGAGCAGATTGCTATCTGTACTCCATTGGACTTCTTTACTACCGT 320  
Db 379 TGGCTTACTTTTGGCAATTAGATTGTTATTTATATTCAGTTGGACTATGGGACTTCCCTGT 438  
Qy 321 TTCTACTATTCCCTGATCTGTGCATCTCAGTTAGCTTCAATGCTTCTTCTCTTATTT 380  
Db 439 CTCTACTATTACTCATTTGCTCATCCCAATTTGGCTTTCAATGCTTCTTCTCTACTT 498  
Qy 381 TCTTAATCACA AAAA ACTTACCCCTATCATTTTAAATCTCTTTTCTCTTAACTATATC 440  
Db 499 CCTCAACTCACTCAAGTTACACACCTTACATAATCAACTCTCTAGTCTCTTCTCACCATTTC 558  
Qy 441 TTCCACCCCTACTTGGCATTCATAATGAG-----GAGACAGACTC 479  
Db 559 TTCAACCCCTCTGTGTTTTCAAAATGAGTCATCATCAGATGATGATGATTCAGATTTC 618

Qy 480 CACAAAAGTTTACAAAAGGAGAGTATGTCAAAGGTTTTCATATGCACCGTTGCTCGCTCTGC 539  
Db 619 CACAAAATCTCCAAAGAAAAGTATGTGATTGGAATTCATATGCACAGTTGGTGCACTCTGC 678  
Qy 540 TGGTTATGGTCTAGTCTTATCCCTACACAGCTAGCTTTCTAAAGAGTCTTAAAGAGCA 599  
Db 679 AGGGTATGGACTATGGCTTTCCCTCACACAGCTTGTGTTCAAAGAGGTCATAAAAAGAGA 738  
Qy 600 AAATTTCTCAGAAAGTTATGGATATGATTAATCTACGTGAGTCTAGTGGCCAGTTGTGTTAG 659  
Db 739 AACATTCAAAGTGGTCTTGGATATGATATATACACTTCCCTTGTGGCTACCCCTGCTAC 798  
Qy 660 CGTGGTGGGCTTTTGTGTAGCAGTGAGTGGAAAACTTTTGAGCAGTGAATAAGTAACATA 719  
Db 799 CTTAGTGGGACTTTTGTCTAGTGGAGAGTGGAGTGCTTTGAAGAATGAAATCAAGAGATA 858  
Qy 720 CAAACATGGGAAGGTATCTCATATTAATGAACCTAGTGTGGACAGCTGTTACCTGGCAGTT 779  
Db 859 TGAGTTGGGGAAGGCTTCATATTTGCTGAACCTCACTTTTCACAGGCATACCTTTGGCAAGT 918  
Qy 780 ATTTCTCCATCGGTGGCACAGGACTGATCTTCGAGCTCTCTCTATTTCTCAAAATGCAAT 839  
Db 919 CTTACCAATGGTGTGTTGGGCGATATTAGTGAGGTTTCTCCCTCTCTCTAAATGCCAT 978  
Qy 840 AAGCGTTTGGGACTCCAGTGGTTCCTATCTTGGCTGTAAATCATTTTCCATGACAAAAAT 899  
Db 979 AAGCGCTTTGGGAGTGCTATTGTTCCAATGCTGCGAGTGTGTTCTTTTCATGACAAAAAT 1038  
Qy 900 GAATGGGTTAAAGGTGATTTCTATGATCCTAGCTATTTGGGGTTTTCACCTTCTATGCTA 959  
Db 1039 GGAATGGCATTAAGGGTATCTCTATGGTGTAGCTATTTGGGGCATTTGTATCATATGTGTA 1098  
Qy 960 CCAACAATATCTTGTATGACAAAACTTGAAGAAAA 994  
Db 1099 TCACAGTACTTGGATGATACAAAGTCCGAAAACA 1133

## RESULT 4

US-10-437-963-82061/c  
; Sequence 82061, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 82061  
; LENGTH: 1513  
; TYPE: DNA  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_81525C.1  
US-10-437-963-82061

Query Match 28.4%; Score 307; DB 17; Length 1513;  
Best Local Similarity 59.4%; Pred. No. 2.5e-76;  
Matches 545; Conservative 0; Mismatches 360; Indels 12; Gaps 1;  
Qy 92 TATACTATGACACGAGGAAACAGTAATCGCTAGCAACGGTAGTTCAACTTGTGCT 151  
Db 1185 TCTACTCAATCAAGGCGGCAATAGCAAGTGATGCCATTCGTCCAAACTGCTGCT 1126

```
Qy 152 TTCC-----TGTCCTACTCCATATATATCTTCTGCTATTTAAACACATGCAA 199
Db 1125 TTCCGATTTTGTTCATTCGCCCTATTTCTTTCCATTTCAAGACATCTTCTACAAACAG 1066
Qy 200 CAACCTAGTAGAGATGGAAAAAGAACCTCACCTAGGAACCGTGTATTTGGTTTACGTAGTGC 259
Db 1065 TCACCTAGTAGTCTGCCCTCAATTTCTATCCCCAAAATTTACTCTGATATATGTGTCC 1006
Qy 260 TTGGACTTCTTGTAGGAGCAGATTTGCTATCTGTACTCCATTTGGACTTCTTTTACTTTACCCG 319
Db 1005 TGGGCTCATCATTTGCTGACACGACTTGTATCTTCCATGGCTACTATATCTTCGG 946
Qy 320 TTCTACCTATTCCTGTATCTGTGATCTGTAGTCTTCAATGCTTTCTTCTTATTT 379
Db 945 TCTCAACATATTCGTCTCATCTGTGTAGTCTGAGCTTGCCTTCAATGCTGTCTTCTCATAT 886
Qy 380 TTCTTAACCTCAAAAACCTTACCCCTCATCTATTTTAAATTTCTTTTCCCTCTTAATATAT 439
Db 885 TCCCTCAATGCTCAAAAATTTCAACCCTCTGATTTTCAATTCGGTAGTCTCTTACGTTTT 826
Qy 440 CTTCCACCTACTTGCATTTCAATTAATAGGAGACAGACTCCACAAAAGTTTCAAAAAGAG 499
Db 825 CTGCTTCACTCTCTTGGAGTTGATGAAGATTTCTCAGGGAACCTACTAGTATATACATGGGA 766
Qy 500 AGTATGTCAAAGGTTTCATATATGCAACCGTTGCTGGCTGCTGGTTATGGTCTAGCTTAT 559
Db 765 AGTACATTTTGGGTTTCTGTGTGACACTAGGGGATCAGCTACATCTCGCTCATCTCT 706
Qy 560 CCCTACACAGCTAGCCCTTCTAAAGTCTTAAAGAACCAAAATTTCTCAGAAGTTATGG 619
Db 705 CCCTGATGCAAGTCACATTTGAGRAGTTATTAGAGGGAGACCTTCTCAGTTGTGTGA 646
Qy 620 ATATGATTAATCTACGTAGTCTAGTGGCCAGTTGTGTAGCGTGGTGGGCTTTTGTCTA 679
Db 645 ACATGCAGATATATACAGCTCTCGTGGCAACATTTGGCTTCTCTTGTGGGTTATTGCAA 586
Qy 680 GCAGTGAGTGAAACTTTTGACGTGAATGGAATGAATTAATAACATGGGAGGTATCCT 739
Db 585 GTGTGTAATGATGACTTTTACAAGGAGAGATGCATGCATTCAGTCTGGGAAGCTGTAT 526
Qy 740 ACATTTAAGACCTAGTGTGGACAGCTGTATCTTGGCAGTTATTTCTCCATCGGTGGCACAG 799
Db 525 ATGTAATGACACTGCTGTGGACGCTATATCTTGGCAGTAGCATCAGTTGGAGTGTGG 466
Qy 800 GACTGATCTTCGAGCTCTCTCTCTATTTCTCAAAATGCAATTAAGCGTTTGGGACTCCAG 859
Db 465 GATTGATCTTTGTGTGTCTATCGCTGTCTTCAAAATGTGATAAGCACCCCTAGCTCTACCCA 406
Qy 860 TGGTTCCCTATCTTGGCTGTATCAATTTTCCATGACAAAATGAATGAGGTAAAGGTGATTT 919
Db 405 TCATTTCTGTTTTTGTCTGTGATTTTCTTTTCATGACAAAGATGGATGGAGTAAAGATTTATAG 346
Qy 920 CTATGATCCTAGCTATTTGGGTTTTTCACTTCTCTATGTCTACCAAAATATCTTGTATGACA 979
Db 345 CTATGCTGATGCCATTTGGGATTTATGTCTATATGGCCACCAATATATATGTTGATGGCA 286
Qy 980 AAACTTGAAGAAAAT 996
Db 285 AGAAAGGTAGAAAGACT 269
```

RESULT 5

```
US-10-424-599-102519
; Sequence 102519, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
```

```
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 102519
; LENGTH: 2214
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_63592C.1
US-10-424-599-102519
```

```
Query Match      28.1%; Score 303.8; DB 16; Length 2214;
Best Local Similarity 60.2%; Pred. No. 2.5e-75;
Matches 540; Conservative 0; Mismatches 352; Indels 5; Gaps 2;

Qy 87 AGGAGTACTATGACAAACGGAGGAAACAGTAATGGCTAGCAACCGTAGTTCACATTTGT 146
Db 476 AAGATTTTATTATGATCAGGGTGGAAATAGTAATAGTAGTCTACTAGTTCAACTGC 535
Qy 147 TGGCTTTCTGTGCTACTTCCATATATATCTTGTGATTTTAAACACATGCAACACTGA 206
Db 536 TGCCTTCCGATCTTGTTCATTCAT---TATTTACAAATTCCTTCACTCCAGAGGCTC 592
Qy 207 TAGAGATGGAAAAAGAAACCTCACCTAGGAACCGGTATTTGGTTTACGTAGTGTGGACT 266
Db 593 AACTTCTGCTTCACTCCCATCAAAATTTCTTTGATATATT--TTGGTCTTTGGAGT 650
Qy 267 TCTTGTAGGAGCAGATTTGTCTATCTGTACTCCATTTGGACTTCTTTTACTTACCGTTTCTAC 326
Db 651 CTTAAATTTGCTGCTGACAAATATGATGACTCCACTGGACTCTTATACCTCTCGGCTTCTAC 710
Qy 327 CTATTCCTGATCTGTGCTACTCAGTTAGCTTCAATGCTTTCTTCTTATTTTCTTAA 386
Db 711 CTATTCCTGATTTGTGCTATCAGATTTAGCTTTTAATGCAAGTCTTCTCATATTTTATCAA 770
Qy 387 CTCAAAAAAATTTACCCCTATCATTTTAAATTTCTTTTCTTCTTAACTATATCTTCCAC 446
Db 771 TTCTCAAAAGTTTCACTGCCCTTGATTAATACTCTACAGTGGTCTCACTTTAICTGTGTC 830
Qy 447 CCTACTTTGCAATCAATAATGAGGAGACAGACTCCAAAAAGTTTACAAAAGAGAGTATGT 506
Db 831 ACTCCTTGTCTTAAACGAAGACACAGATGAACCATCTGGTTTCTCCAAGGAAAGTACAT 890
Qy 507 CAAAGGTTTCATATGACCGTCTGCTGCTGCTGCTGTTATGCTTACTGCTTATCCCTACA 566
Db 891 TATTGGTTTCTATGTACCTTGGAGCTTCTGAGTGTCTGCAAGTACTCTCTTTTGTCTTCCCTCAT 950
Qy 567 ACAGTACGCTTTCTAAAAGTCTTAAAGAACCAAAATTTCTCAGAAAGTTATGATATGAT 626
Db 951 GCAGCTGACCTTTGAGAAAGTTCTGAGAGAGGAAACATTTTCTGTGTTTTGGAAATGCA 1010
Qy 627 AATCTAGCTGAGTCTAGTGGCCAGTTGTGTAGCGTGGTGGGCTTTTTTGTCTAGCAGTGA 686
Db 1011 AATCTACACATCATTCGTTGCCCTCTGGTGTCTCTGTCATAGGCCCTATTGCAAGTGGGA 1070
Qy 687 GTGMAAAACCTTTGAGCAGTGAATGATTAATACTACAAACATGGGAGGATATCCTACATAT 746
Db 1071 ATGGCGTACTTTGCATGGAGAAATGGAGGGTTTTCAGAAAAGGATATGTGCTTATGTTAT 1130
Qy 747 GAACCTAGTGTGGACAGCTGTATTACCTGGCAGTTATTCTCCATCGGTGGCACAGGACTGAT 806
Db 1131 GACTTTGGTTTGGACTTCAATAGCCTGGCAGGTATGCTCTGTGGTGTGTTGTTGGCTTGAT 1190
Qy 807 CTTGAGACTCTCCTCTCTATTTCTCAAAATGCAATTAAGCGTTTTTGGGACTCCCGAGTGTCC 866
Db 1191 CTTCTAGTGTCTTCTCTCTACTCCAATGTTATAAGCACAGTTTCTTTTAGCCGTAACCTCC 1250
Qy 867 TATCTTGGCTGTAATCATTTTCCATGACAAAATGAATGGGTTAAAGGTGATTTCTATGAT 926
Db 1251 TATTGCTGCTGTATAGTTTTTTCATGATAAGATGAATGGGGTGAAGATAAATTTCTATGCT 1310
Qy 927 CCTAGCTATTTGGGGTTTCACTTCTCTATGTCTACCAACAATATCTTGATGACAAAAA 983
```

Db 1311 TTTGGCTCTATGGGTTTTCCTTTATATTTATCAGAATTATCTTGATGATTCAAA 1367

RESULT 6  
US-10-437-963-43860/c  
; Sequence 43860, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 43860  
; LENGTH: 2381  
; TYPE: DNA  
; ORGANISM: Oryza sativa  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(2381)  
; OTHER INFORMATION: unsure at all n locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_46977C.1  
US-10-437-963-43860

Query Match 24.2%; Score 262; DB 17; Length 2381;  
Best Local Similarity 59.1%; Pred. No. 2.1e-63;  
Matches 448; Conservative 0; Mismatches 310; Indels 0; Gaps 0;

QY 235 AACCGTGATTTGGTTTACGTTAGCTGTGGACTTCTTTAGGAGCAGATTGCTATCTGTAC 294  
Db 2039 AAGTCGCGGTGATATACATCGTCTGGGCTCATCATCGCCGACGACATGATGTAC 1980

QY 295 TCATTTGGACTTCTTTACTTACCGTTTACCTTACCTTCCCTGATCTGTCATCTCAGTTA 354  
Db 1979 ACTGTGGGCTCAAGTACCTCCCGCTTCGACCTACTCTGCTCATCTGTCCGACGAGCTC 1920

QY 355 GCGTTCAATGCTTCTTCTTATTTTCTTAATCACAATACTTACCCCTATCATTTTA 414  
Db 1919 GCGTTCAATGCTGCTTCTTCAATACGCTCAACTCCGAGAGGTCACTCTCTGATATTC 1860

QY 415 AATTTCTCTTTTCTCTTAACTATATCTTCCACCTACTTTCATTTCAATAATGAGGAGACA 474  
Db 1859 AACTCGTGTGCTGCTCACCATGTCGCTTGGCTCATCGGAGTCAGAGAGTCTCAG 1800

QY 475 GACTCCAAAAAGTTACAAAAGAGAGATGTCAAAGTTTCAATATGCAACCGTTGCTGG 534  
Db 1799 GGGGTCAACGGGCTCTCGGGAGGAAGTATCTGCTCGTTCGTGTGACGCTGGGGGG 1740

QY 535 TCTGTGTTTATGGTTACTCTTATCCCTACACAGCTAGCTTCTTAAAGTCTTAAG 594  
Db 1739 TGTGCACTTCTGCTGTATCTCGCGCTGATGCACTCACCTCGAGACCATCATCAAG 1680

QY 595 AAGCAAAATTTCTCAGAAGTTATGGATATGATAATCTAGCTGAGTCTAGTGCCAGTTGT 654  
Db 1679 AAGCACACCTTCTCAGCGCTCTCAACATGCGATCTACCGGCTGTGGCGAGGCC 1620

QY 655 GTTAGCGTGGTGGGCTTTTCTAGCAGTGTGAGTGGAAATCTTTGAGCAGTGAATGGAT 714  
Db 1619 GCGTGGTGGTGGGCTGTTCGAGCGCGAGTGGAGTCTGAGGGGGGAGATGAAC 1560

QY 715 AACTACAAACATGGGAGATATCTTACATTTGAACCTAGTGTGACAGCTGTTCCTGG 774  
Db 1559 GCGTTACGTCGGGGGAGTTCTCTCACTGATGACGCTGCTGTGGCGCGCTGTCTGG 1500

QY 775 CAGTTATTCTCCATCGGTGGCACAGGACTGATCTTCGAGCTCTCTCTCTCTATTCTCAAA 834  
Db 1499 CAGGTGGCCAAACATCGGGGTGCTCGGCTCATCTTCGAGGTGCTCGGCGCTCTTCTCAAC 1440

QY 835 GCAATAAGCGTTTGGGACTCCAGTGGTTCCTATCTTGGGCTGTGAATCAATTTTCCATGAC 894  
Db 1439 GTGATCAGCAGCGTGTCTACTGCGGTCATCCCGTTCCTTCGCGGTGCTGTGTTCCACGAC 1380

QY 895 AAAATGAATGGGTTAAAGTGATTTCTATGATCCTAGCTATTTGGGTTTCACTTCCTAT 954  
Db 1379 AGGATGAACGGGTGAAGATCGTGCCATGCTGATTCGAATTTGGGATTTATTCGTAT 1320

QY 955 GTCTACCAACATATCTTGTGACAAAAAATTGAAGAA 992  
Db 1319 CTGTTTCAGCACTATCTAGTGGCAAGAAAGCAAGAA 1282

RESULT 7  
US-10-416-898-9  
; Sequence 9, Application US/10416898  
; Publication No. US20040172670A1  
; GENERAL INFORMATION:  
; APPLICANT: Yale University  
; APPLICANT: Walker, Elisabeth  
; APPLICANT: Dellaporta, Stephen  
; TITLE OF INVENTION: MAIZE YELLOW STRIPE AND RELATED GENES  
; FILE REFERENCE: 44574-5106-US  
; CURRENT APPLICATION NUMBER: US/10/416,898  
; PRIOR FILING DATE: 2003-05-16  
; PRIOR APPLICATION NUMBER: PCT/US01/43101  
; PRIOR FILING DATE: 2001-11-16  
; PRIOR APPLICATION NUMBER: US 60/249,222  
; PRIOR FILING DATE: 2000-11-16  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 9  
; LENGTH: 83698  
; TYPE: DNA  
; ORGANISM: Zea mays  
US-10-416-898-9

Query Match 21.7%; Score 234.2; DB 17; Length 83698;  
Best Local Similarity 54.3%; Pred. No. 1.3e-54;  
Matches 495; Conservative 0; Mismatches 413; Indels 3; Gaps 1;

QY 92 TATACATGACACCGGAGAAACAGTAAATGGCTAGCAACGGTAGTCTCAACTGTGTGGCT 151  
Db 67301 TTTATTACAATGAAGGTGGAACAGTAATGGATCTCTACTCTTGTTCAAACTGGTGGCT 67360

QY 152 TTCTGTGCTACTTCCATATATATATCTTGTGATTTTAAACACATGCAACAACATGATAG 211  
Db 67361 TTCCATTTCTTTATCTCCCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 67417

QY 212 ATGAAAAAAGAACCTCACCTAGGAACCGTGTATTGGTTTACGTAGTGTCTGAGCTTCTTG 271  
Db 67418 CT 67477

QY 272 TAGGAGCAGATGTCTATCTGTACTCCATGTGACTTCTTTACTTACCCGTTTCTACCTATT 331  
Db 67478 TTGGTTTAGATAAATTTTTTATCTCTGTGGACTTTGTATCTCTCTCTCTCTCTCTCT 67537

QY 332 CCCTGATCTGTGCATCTCAGTTAGCTTCAATGCTTCTCTCTCTCTCTCTCTCTCTCTCTCT 391  
Db 67538 CGATTCTATGTCTTACAGTTAGCTTCAATGGTGTCTCTCTCTCTCTCTCTCTCTCTCT 67597

QY 392 AAAAATTTACCCCTATCATTTTAAATCTCTTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 451  
Db 67598 AGAAATACATTTGTTGATTTCTCTCAGTGTGTTCTCTCTCTCTCTCTCTCTCTCTCTCT 67657

QY 452 TTGCATTCATTAATGAGGAGAGACAGACTCCA AAAAGTTTCAAAAAGGAGAGATGTGCAAG 511  
Db 67658 TTTCTCTTGAGCATGATTCAAAATAGCCCATCAGGAGATTTCTAAGTGGAGTACTTGAATTG 67717

QY 512 GTTTCATATGACCGTTGCTGGTCTGCTGGTCTGCTAGTCTTAGTCTTATCCCTCAACAGC 571  
Db 67718 GGTGTTCTGTGACAGTTTGTGCTCTCTTATCTATCTCTTCTAGCTCTCTCTTATGACGT 67777  
QY 572 TAGCCTTTCTAAAGTCTCTAAAGACGAAATTTCTCAGAAAGTTATGATATGATAATCT 631  
Db 67778 TTTCTTTTCGAGAGGTTCTCAGAGTGTAGATCTCTCTATGTTCTCGAGATGCANATCT 67837  
QY 632 AGCTGAGTCTAGTGGCCAGTTGTGTTAGCGTGGTGGGCTTTTGTCTAGCAGTGTAGTGA 691  
Db 67838 ATACGTCGCTTCTGCTTCTTGTGTAGCGGTTATCGGATTTGTGCAAGCGGGAAATGA 67897  
QY 692 AAATCTTCAGCAGTGAATGATTAACATCAACATGGAAGGTATCTTACATTTATGACCT 751  
Db 67898 TGTGTTGAGTGTGGAGATGGAAGTTTTCAGGAAGGTCAAGTCAATTTATGTTTGTGACTT 67957  
QY 752 TAGTGTGACAGCTCTTACCTGGCAGTTTATTTCTCCATCGGTGGCAGCAGTGTATCTTCG 811  
Db 67958 TGGTCGGGGCAGCGGTTTCGTGTCAATTTGGGTGTGTAGGAGCCGTGTCTTATTTTC 68017  
QY 812 AGCTCTCTCTATCTCAATGCAATGAAGCGTTTGGGACTCCAGTGTGTTCTTATCT 871  
Db 68018 TGGTCTCTGCTGTTTCAAACTTATTAGTAGCTCTCACTCATTTGTTACGCTCTCTCG 68077  
QY 872 TGGCTGTAAATCATTTTCCATGACAAATGAATGGTTAAAGGTGATTTCTATGATCCTAG 931  
Db 68078 CGGCCATTTGCGGTGTTCATGACAGCTGACTGAGGTTAAGATGGTGGCATGCCATCG 68137  
QY 932 CTATTGCGGTTTCACTTCTCTATGCTACCAACATATCTTGTGACAAAACTTGAAGA 991  
Db 68138 CTTTCACAGGATTCAGTTTATATCTACCAAGACTATCTTGATGACTTGAAGTACAAA 68197  
QY 992 AAAATCATGAA 1002  
Db 68198 GAGCAGAGAA 68208

RESULT 8  
US-10-424-599-75161  
; Sequence 75161, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J  
; APPLICANT: Kovalic, David K  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 75161  
; LENGTH: 997  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_38884C.1  
US-10-424-599-75161

Query Match 20.7%; Score 223.4; DB 16; Length 997;  
Best Local Similarity 58.4%; Pred. No. 1.2e-52;  
Matches 464; Conservative 0; Mismatches 301; Indels 30; Gaps 3;

QY 87 AGGAGTATATGATCAACGGAGGAAACAGTAAATGGCTAGCAACGGTAGTTCAACTTGT 146  
Db 202 AAGATTATATGACATGGTGTAAATGCAATGGATGGCAACATTTGTTCAATCAGC 261  
QY 147 TGGCTTTCTGTGCTACTTCCA---TATTATATCTGTTCATTTAAACACATGCAACAC 203  
Db 262 TGGATTCCCTGTACTTCTTCCACTCTCTTTTACTTCCCAAGACAAACACATGCCAAGTT 321

QY 204 TGATAGAGATGG-----AAAAAGAACTTCACCTAGGAACCG 239  
Db 322 TAATAACAATCCAGCAATAATGATATTCTCAAAACAAACCAAAATTTCTCCACCT 381  
QY 240 TGTAATGTTTACGTAGTGTGGACTTCTGTAGGAGCAGATTTGTCTATCTGTACTCCAT 299  
Db 382 TGTCTTCTCTATCTAGCTTTTGGCCTAAATCTCACAGGGGCAACCTTGATGTATTCTTA 441  
QY 300 TGGACTTCTTACTTACCCCGTTTCTACTATTTCCCTGATCTGTGCATCTCAGTTAGCCTT 359  
Db 442 TGGACTTTTATATCTCCCTCTTCCACCTATTTCTACTATGTGCCCCCAATTTAGTCTT 501  
QY 360 CAATGCTTTCTCTCTTATTTTCTTAACACAAAACTTTACCCCTATCAATTTTAAATTC 419  
Db 502 CAACGCGTGTCTCTTCTCTCAATTTCCAGAAATTCACAGCATTCATATCAATTC 561  
QY 420 TCTTTTCTCTTAATAATATCTTCCACCTTCTGCAATTAATAATGA---GGAGACAGA 476  
Db 562 TGTAGTCTCTTACCATAATCAGCTTCTCTGCTTGAATCAACTCCGATTTCCGACGAGGA 621  
QY 477 CTCCACAAAAAGTTACAAAAGGAGATGTCAAAGGTTTTCATATGACACCGTTGCTGCTC 536  
Db 622 CTCAACCGGCTTTTCCAGAGAAAGCATGTAATTGGATTTCTTGCACCATTTGGTGCATC 681  
QY 537 TGCTGGTTATGCTAGTCTTATCCCTCAACACAGCTAGCCCTTTCTAAAAGTCTCTAAAGAA 596  
Db 682 AGCTACATTTCTTTGTACCTCTCTTGTGACGCTTTCTTTTCAGAAAGTTATAAGAG 741  
QY 597 GCAAAATTTCTCAGAAAGTTATGATATGATAATCTACGTGAGTCTAGTGGCAGTTGTGT 656  
Db 742 AGAAACCTTTTCTGCTGTGTGGACATGCAATTTTACCCATCTCTTCTATGCTACATGTC 801  
QY 657 TAGCGTGTGGGCTTTTGTCTAGCAGTGTGAGTGGAAACTTTTGACAGCTGAAATGGATA 716  
Db 802 TTGTGTAGTAGGGTTGTTGCAAGTGGAGAGTGGAAAGTTTGAACATGAGATGAAGGG 861  
QY 717 CTACAAACATGGGAAGGTATCTACATTAATGAACCTAGTGTGACAGCTGTACTGCGCA 776  
Db 862 ATATGATAAGGAAGTGTGCCGTTGTAATGACTGTACTATAGATTCTGTGACATGGTA 921  
QY 777 GTTATTTCTCATCGTGGGACAGCAGTGTATCTTGCAGCTCTCTCTCTATTTCTCAAATGC 836  
Db 922 TATATGCTCACTAGGATGCTAGGGGGAGCTTTGGAGGAGTGTGCAATTTGTTGCGGATGT 981  
QY 837 AATAAGCGTTTGGG 851  
Db 982 GAGAAAGTACCTGGGG 996

## RESULT 9

US-10-767-701-9000  
; Sequence 9000, Application US/10767701  
; Publication No. US20040172684A1  
; GENERAL INFORMATION:  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53535)B  
; CURRENT APPLICATION NUMBER: US/10/767,701  
; CURRENT FILING DATE: 2004-01-29  
; NUMBER OF SEQ ID NOS: 63128  
; SEQ ID NO 9000  
; LENGTH: 728  
; TYPE: DNA  
; ORGANISM: Sorghum bicolor  
; FEATURE:  
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS69833\_1  
US-10-767-701-9000

Query Match 16.6%; Score 179.2; DB 17; Length 728;  
Best Local Similarity 61.2%; Pred. No. 3.9e-40;

Matches 289; Conservative 0; Mismatches 183; Indels 0; Gaps 0;

QY 532 GCCTGCTGCTGTTATGCTTAGTCTTATCCCTACACAGCTAGCCTTTCTAAAGTCTTAA 591  
Db 4 GCCTCAGCCACATACCTACTTAATCTCTCTAATGCAAGTCACATTTGAGAGGTTATT 63  
QY 592 AAGAAGCAAAATTTCTCAGAAATTAATGATATGATATATCTACGTGAGTCTAGTGGCAGT 651  
Db 64 AAGAAGCAAAACGTTCTCAGTCTGTTGAATATGAGATATATACAGCACTAGTGGCAACA 123  
QY 652 TGTGTTAGCTGCTGCTGGGCTTTTGTAGCAGTGGTGAAGAACTTTGACGAGTGAATG 711  
Db 124 ATAGCTCTCTGTTGGTTATTGCAAGCGGTGAGTGAAGAACTTTAGAGGAGAGATG 183  
QY 712 GATAACTACAAACATGGGAAGGTATCTCAATATGAACCTAGTGTGGACAGCTGTTACC 771  
Db 184 CATGCCCTTCAGCTCAGGAGGTTGTATATGTGATGACACTTCTATGGAGCTGCTGTTAT 243  
QY 772 TGGCAGTTATTCATCGGTGGACAGAGCTGATCTTCGAGCTCTCCTCTCTPAITCTCA 831  
Db 244 TGGCAGATAGCTTCTGTAGAGTGGTGGGTTGATCTTTGTGTGTCATCACTCTTTTCA 303  
QY 832 AATGCAATAGCGTTTGGGACCTCCAGTGGTTCCTATCTTGGCTGTAATCAATTTCCAT 891  
Db 304 AATGTGATAGCACACTAGCTCTACCCATCAITCCAGTTTTTGTGTGATTTTCTTCCAC 363  
QY 892 GACAAATGAATGGGTTAAAGGTGATTTCTATGATCTAGCTATTTGGGGTTTCACTCC 951  
Db 364 GACAAGTGAATGGAGTGAAGATTTATGTTATGTTGATGGCCATCTGGGATTTGTTCA 423  
QY 952 TATGCTACCAACAATATCTTGTATGACAAAACTTGAAGAAAATCATGAAA 1003  
Db 424 TATGGATACCAATATATATATAGTACAGAAAGCTTAGGAAGACTTCAGTAA 475

RESULT 10  
US-10-424-599-114931  
; Sequence 114931, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J  
; APPLICANT: Kovalic, David K  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 114931  
; LENGTH: 522  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_74793C.1  
US-10-424-599-114931

Query Match 13.8%; Score 149; DB 16; Length 522;  
Best Local Similarity 56.8%; Pred. No. 1.3e-31;  
Matches 295; Conservative 0; Mismatches 225; Indels 1; Gaps 1;

QY 353 TAGCCCTTCAATGCTTCTCTTATTTCTTAACCTCACAAAACCTTACCCCTATCATTT 412  
Db 1 TAAATTTGAATGAGTGCTCACTTTCTTCATCAATTTCCAAAGATTCATGCAATGATCT 60  
QY 413 TAAATTTCTTTTCTTTAACTATATCTTCCACCCTACTTGCATTTCAATAATAGGAGA 472  
Db 61 TAAACTCTAATCGTCTTACCATATCAGTTACCTGATTTGCACTCAATAGTATCTG 120  
QY 473 CAGACTCCAAAGTTACAAAGGAGATGTCTCAAGGTTTCAATATGCAACGTTGCTG 532  
Db 121 AGGAAACAAAGAACCTTTCCAAACAGAGCAAAATTAATTTGGGTTCTTCTGTGCCCTAGGTG 180

QY 533 CGTCTGCTGTTATGCTTAGTCTTATCCCTACACAGCTAGCCTTTCTAAAGTCTTAA 592  
Db 181 CATCTGCCATATTTGCATTTGCATCACTCTCTTATGCAATTTATTTTGAGAAAATTTATA 240  
QY 593 AAGAAGCAAAATTTCTCAGAAATTAATGATATGATATATCTACGTGAGTCTAGTGGCAGTT 652  
Db 241 AGACTGAAACTTTTCTACTGTTATTAAGCATGATTTTTTACCCTAATGATCGTGGGTACAA 300  
QY 653 GTGTTAGCTGCTGCTGGGCTTTTGTAGCAGTGGTGAAGAACTTTGACGAGTGAATG 712  
Db 301 TTGTTGGCTCTGTTGGATTTGCTTGAAGTGGAGATTTGGAGAACTATGGGAATGGAGATGA 360  
QY 713 ATAACCTACAAACATGGGAAGGTATCTCAATATGAACCTAGTGTGGACAGCTGTTACCCT 772  
Db 361 AGGAATTCGAAACGAGTAGTGTGATATGATGATGACTCTGTTGTACTTCTGTGACAT 420  
QY 773 GCGAGTTATTTCCATCGGTGGACAGAGCTGATCTTCGAGCTCTCCTCTCTATTTCTCA 831  
Db 421 GCGAGATAGTGTGTGTTGATGCTGGGTTGATTTTTTGAGGTATCTTCAATACCCTCAG 480  
QY 832 AATGCAATAGCGTTTGGGACCTCCAGTGGTTCCTATCTT 872  
Db 481 CCTGCATAAGCAATCTGGAATTGAAACATAGCTCTCTATCTCT 521

RESULT 11  
US-10-425-114-12485  
; Sequence 12485, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 12485  
; LENGTH: 1386  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: 701179255\_FLI  
US-10-425-114-12485

Query Match 13.1%; Score 141.6; DB 16; Length 1386;  
Best Local Similarity 50.5%; Pred. No. 2.9e-29;  
Matches 370; Conservative 0; Mismatches 359; Indels 3; Gaps 1;

QY 248 TTTACGTAGTGTCTGGACTTCTTTGAGGAGCAGATTTGCTATCTGTACTCTCATTTGGACTTC 307  
Db 411 TCTACGCGGCATCGGGTCAATGCTGCTTCGACAACTCATGTACTCTGCTACGCGCTGC 470  
QY 308 TTTACTTACCGGTTTCTACCTATTTCCCTGATCTGTGCACTCTCAGTTAGCCTTTCAATGCTT 367  
Db 471 AGTACTCGCGGTGTCACCTTCTCGTCTGGCGCGACGAGCTGGGCTTCAACGCCA 530  
QY 368 TCTTCTCTTATTTCTTAACCTCACAAAACCTTACCCCTATCATTTTAATTTCTCTTTTCC 427  
Db 531 TCACCTCGCGCTCATCAACGCGAGCGGTTTTCACGCGCTGATCGCCAACTCCGTGGTGG 590  
QY 428 TCTTAACATATATTTTCCACCTTACTTTGATTTCAATTAATGAGGAGACAGA---CTCCACAA 484  
Db 591 TGCTCACCTTCTTCGCGCGGATCTTCGCGGTGGATTCCTCTCCGACGAGACCTCCAGA 650  
QY 485 AAGTTACAAAAGGAGAGTATGTCAAGGTTTTCATATGCAACGCTTCTCGCTCGCTGCTGTT 544





Db 895 TCACCTCGCGCTCATCAACGCGCAGCGGTTACGGCGCTGATCGCCAACTCCGTGGTGG 954  
Qy 428 TCTTAACATATCTTCACCCCTACTTGCATTTCAATATGAGGACAGAGA---CTCCACAA 484  
Db 955 TGCTCACCTTCTCGCGCGGATCTCGCGCTCGGATCCTCCTCGAGACCTCCAGCA 1014  
Qy 485 AAGTTACAAAGAGAGATATGTCMAAGGTTTCATATGACACCGTTCGTGCTGCTGCTGGTT 544  
Db 1015 ACGTGCGCGCGCAAGTACCGCTGGGGTTCGTCTCTGCTGCTGCGCGCTCGCGGCTGT 1074  
Qy 545 ATGCTCTAGTCTTATCCCTCAACACAGCTAGCTTTCTTAAAGTCTTAAAGAACAAATTT 604  
Db 1075 TCGGCTCATTTTGTCTCTCTTTCAGCTCACCTTCGAGAGGGTGCAGGGTCGGACGG 1134  
Qy 605 TCTCAGAAGTTATGGAATGATATCTACGTGAGTCTAGTGGCCAGTTGTTAGGGTGG 664  
Db 1135 CGCGGTGGGTGCTCGGATGCAGATGCACACCAACCTGGTGGTGGCGGTGCTCGGTGG 1194  
Qy 665 TGGGGCTTTTGTCTAGCAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 724  
Db 1195 CGGGGTGCTCGGCTCGGGGAGTGGCGGACGATCCGGGGGAGATGGCGTCTTCAAAG 1254  
Qy 725 ATGGAAGGTATCTCATATATGAACCTAGTGTGGACAGCTGTACCTGGCAGTTATCT 784  
Db 1255 ACGGAGGACGAGTACGTCTGACGCTGGTGGGACGCGGGTGTGGGACGAGCGG 1314  
Qy 785 CCATCGGTGGCAGCAGCACTGATCTTCGAGCTCTCTCTTATTTCTCAAAATGCAATAGCG 844  
Db 1315 CCGTGGCTTGGTGGCTGATCATGAGGGTGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1374  
Qy 845 TTTTGGGACTCCAGTGGTTCCTATCTTGGCTGTATCATTTTCCATGACAAATGAATG 904  
Db 1375 CGTGGCGCTGCGGCTGGTCCAGTGTTCGGGTGGCGCTGTTCGGGGGACGAGTACCG 1434  
Qy 905 GGTAAAGGTGATTTCTATGATCCTAGCTATTTGGGGTTTCACTTCTATGTTCTACCAAC 964  
Db 1435 GTATAAGATCGTGGCCATCTCATGCGGTACGGGGTTCCTCTCTCATATGATACCAGC 1494  
Qy 965 AATATCTTGATG 976  
Db 1495 ACTACATCGACG 1506

RESULT 14  
US-10-424-599-30274  
; Sequence 30274, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 30274  
; LENGTH: 1233  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_12733C.1  
US-10-424-599-30274  
Query Match 12.2%; Score 131.8; DB 16; Length 1233;  
Best Local Similarity 47.3%; Pred. No. 1.7e-26;  
Matches 431; Conservative 0; Mismatches 477; Indels 3; Gaps 1;  
Qy 92 TATACATGACACGAGGAAACAGTAATGGCTAGCAGCGTAGTTCAACTGTTGGCT 151  
Db 8 TCTACTTCAATCCAGGTGGCAACGAATCTGGCTCTCTAGCTTCCTTGAACACTGCAGCCA 67

Qy 152 TTCTGTGTCTACTTCCATATATATATCTGTGCTATTTAAACACATGCAACAACTGATAGAG 211  
Db 68 ATTGCGCTCGAGTCTTCCATAATCATTTCTTACATTATATAACGACGCTCGTATCGCT 127  
Qy 212 ATGGAAAAAGAACCTCACCCTAGAACCGTGTAT---TGGTTACGTTAGTGTCTTGGACTTC 268  
Db 128 CTTTCGCGCTTAATCTCTATCAAGCCCTCTCTCTCCGCTCCGCGCTCATCGGACTCC 187  
Qy 269 TTGTAGGACAGATTGCTATCTGTAATCTCCATTTGGACTTCTTACTTACCGTTCCTACCT 328  
Db 188 TCACCGGCTCGACGACTACCTCTAGCCTCGCGGTGGCTCGCTTCCGGTCTCCACT 247  
Qy 329 ATTCCCTGATCTGTGATCTCAGTTTAGCCTTCAATGCTTCTTCTTATTTTCTTAACT 388  
Db 248 TCTCTCTAATCAAGCTCCACCTCCGCTTCAACGCGCTTTCGCGCTTCCCTCGTCC 307  
Qy 389 CACAAAACCTTACCCCTATATTTAAATCTCTTTTCTCTTAACTATATCTTCCACCC 448  
Db 308 GCCACAGGTTTCAGCGCTTACTCCGTCACCTCGGCTGTCTTCTCACCGTCCGCTGTGG 367  
Qy 449 TACTTGCATTCATTAATGAGGAGACAGACTCCACAAAAGTTACAAAAGGAGATATGTC 508  
Db 368 TTCTGCTCTGCGTTCACGCGGACCGCCCGCGGTGAGTCGAGTCGCCAGTACGTGA 427  
Qy 509 AAGGTTTTCATATGCAACCGTGTGCTGTGCTGTGTTATGTTCTAGTCTTATCCCTACAAC 568  
Db 428 TTGTTTGTGTTATGATCTTTCGCGCTCGCGCTGTATGGGTTGTTTTCGCGTGTATGG 487  
Qy 569 AGCTAGCCTTTCTAAAGTCTTAAAGAGCAAAATTTCTCAGAAAGTTATGGATATGATA 628  
Db 488 AGTGGTGTACAAAAGAGAGCAGCGATATCACGCTACTCTCTGGTTCATGGAGATTCAGC 547  
Qy 629 TCTAGTGTAGTCTAGTGGCAGTGTGTTAGCGTGTGGGGTCTTTTGTCTAGCAGTACT 688  
Db 548 TTGCTGTGTGCTTCTTGTCTACTTATCTGACCGTGTGGAATGATAATCAATAATGACT 607  
Qy 689 GGAACCTTTTGAGCAGTGAATGGATACTACAAACATGGGAAGGTATCTCATTTATGA 748  
Db 608 TCAAGGTGATTCGCGAGAGCAAGAGATTTTAAGCTTGGGGAACAAAGTACTACGTTG 667  
Qy 749 ACCTAGTGTGACAGCTGTTTACTCGCAGTTATTTCCATCGGTGGCAGCAGGACTGAT 808  
Db 668 TGTGGTGTGGAGTGAATATGTGCGAGTTTTTCTTCTGGGAGCAATAGGGGTATCT 727  
Qy 809 TCGAGCTCTCTCTCTATTTCAAAATGCAATGAGGCTTTTGGGACTCCGAGTGTCTCTA 868  
Db 728 TTTGTGCTCTGCTTCTTGTGTTGCTGTTATTAATTTGCTGCTTTTCTTCCAGTGACGGAAG 787  
Qy 869 TCTTGGCTGTATCAATTTTCCATGACAAAATGAATGGGTAAAGGTGATTTCTATGATCC 928  
Db 788 TTTTGGCTGTATTTCTATACAAAGAGAGCTTTCATGACAGAAAGGGGTGCTTTGGTGC 847  
Qy 929 TAGCTATTTGGGGTTTCACTTCTATGCTACCAACAATATCTTGTATGACAAAACCTGA 988  
Db 848 TCTCTCTTTGGGGTTTGTGCTCTATTTCTATGGAGAGATAAAACAAGACAGGAAAGA 907  
Qy 989 AGAAAAATCAT 999  
Db 908 ACAAGAAATCGT 918

RESULT 15  
US-10-425-114-10516  
; Sequence 10516, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei



GenCore version S.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 1, 2004, 05:48:03 ; Search time 99.7262 Seconds  
(without alignments)  
7704.726 Million cell updates/sec

Title: US-09-913-767-6  
Perfect score: 1081  
Sequence: 1 caaatccaacagtccaagat.....ataagctgattttgaaag 1081

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA:\*  
1: /cgn2\_6/prodata/1/ina/5A COMB.seq:\*  
2: /cgn2\_6/prodata/1/ina/5B COMB.seq:\*  
3: /cgn2\_6/prodata/1/ina/6A COMB.seq:\*  
4: /cgn2\_6/prodata/1/ina/6B COMB.seq:\*  
5: /cgn2\_6/prodata/1/ina/PCUS COMB.seq:\*  
6: /cgn2\_6/prodata/1/ina/backfileseq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	57.8	5.3	7218	1	US-08-232-463-14
2	46.8	4.3	1141	4	US-09-806-708B-22
3	42.2	3.9	832	4	US-09-621-976-2813
4	41.2	3.8	1141	4	US-09-806-708B-22
5	40.6	3.8	72604	3	US-09-268-992-7
6	40.6	3.8	72604	3	US-09-657-474-7
7	40.4	3.7	1055	4	US-09-806-708B-23
8	39.2	3.6	474	4	US-09-621-976-18033
9	38.6	3.6	2674	3	US-09-019-095A-1
10	37.8	3.5	1664976	4	US-08-916-421B-1
11	37.8	3.5	1664976	4	US-09-692-570-1
12	36.8	3.4	248	3	US-09-007-005-32
13	36.8	3.4	248	3	US-09-244-796-32
14	36.8	3.4	277	3	US-09-007-005-3
15	36.8	3.4	277	3	US-09-244-796-3
16	36.8	3.4	8093	4	US-10-204-708-31
17	36.6	3.4	1536	4	US-09-107-532A-2458
18	36.2	3.3	4726	2	US-08-633-770A-11
19	35.8	3.3	70000	4	US-09-851-896-3
20	35.6	3.3	24358	4	US-09-392-812A-1
21	35.4	3.3	705	4	US-09-270-767-5061
22	35.4	3.3	705	4	US-09-270-767-20343
23	35.4	3.3	909	4	US-09-710-279-3061
24	35.4	3.3	948	3	US-09-134-001C-1684
25	35.4	3.3	3554	4	US-09-710-279-4291
26	35	3.2	5217	4	US-09-100-703A-25
27	35	3.2	5230	4	US-09-100-703A-26

C 28	35	3.2	5230	4	US-09-838-718A-5	Sequence 5, Appli
C 29	35	3.2	5231	4	US-09-100-703A-27	Sequence 27, Appl
C 30	35	3.2	5770	4	US-09-838-718A-7	Sequence 7, Appli
C 31	35	3.2	5870	4	US-09-838-718A-8	Sequence 8, Appli
C 32	35	3.2	5906	4	US-09-838-718A-6	Sequence 6, Appli
C 33	34.8	3.2	548	4	US-09-621-976-894	Sequence 894, App
C 34	34.8	3.2	696	3	US-09-461-697-193	Sequence 193, App
C 35	34.8	3.2	699	3	US-09-461-697-191	Sequence 191, App
C 36	34.8	3.2	717	3	US-09-461-697-189	Sequence 189, App
C 37	34.8	3.2	774	3	US-09-461-697-187	Sequence 187, App
C 38	34.8	3.2	819	3	US-09-461-697-185	Sequence 185, App
C 39	34.8	3.2	1669	3	US-09-461-697-184	Sequence 184, App
C 40	34.6	3.2	832	4	US-09-621-976-2813	Sequence 2813, Ap
C 41	34.6	3.2	2319	4	US-09-561-077C-15	Sequence 15, Appl
C 42	34.6	3.2	2319	4	US-09-221-014-15	Sequence 15, Appl
C 43	34.6	3.2	3551	4	US-09-561-077C-16	Sequence 16, Appl
C 44	34.6	3.2	3551	4	US-09-221-014-16	Sequence 16, Appl
C 45	34.6	3.2	7113	4	US-09-561-077C-26	Sequence 26, Appl

ALIGNMENTS

RESULT 1  
US-08-232-463-14  
; Sequence 14, Application US/08232463  
; Patent No. 5670367  
; GENERAL INFORMATION:  
; APPLICANT: DORNER, F.  
; APPLICANT: SCHEIFLINGER, F.  
; APPLICANT: FALKNER, F. G.  
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSES: Foley & Lardner  
; STREET: 1800 Diagonal Road, Suite 500  
; CITY: Alexandria  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22313-0299  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/232,463  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/935,313  
; FILING DATE:  
; APPLICATION NUMBER: EP 91 114 300.6  
; FILING DATE: 26-AUG-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BENT, Stephen A.  
; REGISTRATION NUMBER: 29,768  
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703)836-9300  
; TELEFAX: (703)683-4109  
; TELEX: 899149  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7218 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; CLONE: pTZ9pt-Fls  
US-08-232-463-14

Query Match 5.3%; Score 57.8; DB 1; Length 7218;





```
US-09-657-474-7
; Sequence 7, Application US/09657474
; Patent No. 6399762
; GENERAL INFORMATION:
; APPLICANT: Chen, H.
; APPLICANT: Freimer, N.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
; TITLE OF INVENTION: METHODS AND TREATING CHROMOSOME-18p RELATED DISORDERS
; FILE REFERENCE: 7853-138
; CURRENT APPLICATION NUMBER: US/09/657,474
; CURRENT FILING DATE: 2000-09-07
; PRIOR APPLICATION NUMBER: 09/268,992
; PRIOR FILING DATE: 1999-03-16
; PRIOR APPLICATION NUMBER: 09/236,134
; PRIOR FILING DATE: 1999-01-22
; PRIOR APPLICATION NUMBER: 60/106,056
; PRIOR FILING DATE: 1998-10-28
; PRIOR APPLICATION NUMBER: 60/088,312
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/078,044
; PRIOR FILING DATE: 1998-03-16
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 72604
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified base
; LOCATION: all n positions
; OTHER INFORMATION: n=a, c, g, or t
US-09-657-474-7

Query Match      3.8%; Score 40.6; DB 3; Length 72604;
Best Local Similarity 54.3%; Pred. No. 0.31;
Matches 82; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

QY 288 TCTGACTCCATGGACTCTTACTACCGTTTCTACCTATTCCTGATCTGTCGATC 347
DB 33298 TTGGAATCCATCTGATTCCTCTATTTTCCAGTTGATGCCATGCTCTTGGCCATC 33357

QY 348 TCAGTATGCTTCAATGCTTCTCTCTATTTTCTTAACACACAAACTTACCCTAT 407
DB 33358 TTCTTAATTAAGATGCTTTCCTCTAAATCTCTTAATTAATCAAGCAGCTATCTACA 33417

QY 408 CATTTTAAATCTCTTTTCTCTTAACATA 438
DB 33418 TATTTTGTAATCCCTTAAATCTTGAGCATA 33448

RESULT 7
US-09-806-708B-23/c
; Sequence 23, Application US/09806708B
; Patent No. 6784342
; GENERAL INFORMATION:
; APPLICANT: The University of British Columbia
; TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
; FILE REFERENCE: 4810-58741
; CURRENT APPLICATION NUMBER: US/09/806,708B
; CURRENT FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/147,133
; PRIOR FILING DATE: 1999-08-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 23
; LENGTH: 1055
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (1)..(1055)
; OTHER INFORMATION: consensus sequence of A.t. and L.a. FAEI promoters
US-09-806-708B-23
```

```
Query Match      3.7%; Score 40.4; DB 4; Length 1055;
Best Local Similarity 21.5%; Pred. No. 0.035;
Matches 74; Conservative 99; Mismatches 169; Indels 2; Gaps 1;

QY 150 CTTTCCTGTGCTACTCCATATATATCTGTCAATTTAAACACATCAACAACCTGATAG 209
DB 415 MTTAANNKAGAMWTTTAAAGNNTWTAATAATMAAACAARAYCCNNNWAACAMAKWMA 356

QY 210 AGATGGAAGAAACCTCACCTAGGAACCGTGTATTTGGTTTACGTAGTGTGACTTCT 269
DB 355 WTKWYAGAACNNNNKTAMTCRRRAWMYSAWTTTMAAATSWWKWYTTTTRKTTMAAANN 296

QY 270 TGTAGGAGCAGATTGCTATCTGTACTCCATGGACTCTTTACTTACCCTGTTCTACTA 329
DB 295 NNNNNWAKCKTTSAMWAMWMTWCTCGARTWGGATYMAAACCTAAGAGNSMTYW--CWA 238

QY 330 TTCCCTGATCTGTGCATCTCAGTTAGCTTCAATGCTTCTCTCTTCTTATTTCTTAACTC 389
DB 237 ATYMGTTTNNNNNNKAWTTTCTMAWCAYTTWTARTYKRYTYCYATYTYTSMYMGKM 178

QY 390 AAAAAAATACCCCTATCATTTTAAATCTCTTTTCTCTTCTTAACTATATCTTCCACCT 449
DB 177 WYWARAAAYASTNCTSTCRWKTARGWYVAMRYTWAASGNCNNNAKWCAMWYACATN 118

QY 450 ACTTGCAATCAATAATGAGGAGACAGACTCCACAAAAAGTTACAA 493
DB 117 NNAMAWYACAYWMTYTAGKAAWTTNNKTASGKWMYAMMKTTWYWA 74

RESULT 8
US-09-621-976-18033/c
; Sequence 18033, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano J.V.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 18033
; LENGTH: 474
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 16
; OTHER INFORMATION: n=a, g, c or t
US-09-621-976-18033

Query Match      3.6%; Score 39.2; DB 4; Length 474;
Best Local Similarity 14.3%; Pred. No. 0.052;
Matches 44; Conservative 127; Mismatches 136; Indels 0; Gaps 0;

QY 591 AAAGAAGCAAAATTTCTCAGAAGTTATGGATATGATAATCTACGTAGTCTAGTGCCAG 650
DB 307 ARRTWARRAAMYTWITWYMKSGCMWKMWRKYMKYKCCYWWWRKYCCMRKKRRRG 248

QY 651 TTGTGTTAGCGTGTGGGCTTTTGTCTAGCAGTGCAGTGGAAACTTTTGACGACTGAAT 710
DB 247 KKRARSYKSGYKGSYTYKGGKGGTGKGAWSKKRRRAMTTTTTYKKKWRSYWAA 188

QY 711 GGATAACTACAAACATGGGAAGGTATCTCAATTAATGAACTAGTGTGGACAGCTGTAC 770
DB 187 RGKTTKKWAAMMMCMWAWMYRSMYYKKWAAWKKCCCMARKKGGSSGRWYWKRRGT 128

QY 771 CTGCAGTATTTCTCCATCGGTGCACAGGACTGATCTTCGAGCTCTCTCTATTCTC 830
DB 127 CMRRRRWYRRRCMWYCTKGSMAWMSGKKRAAMCCSKYVYTWCMMAAAWWTWAA 68
```



```
; NAME/KEY: misc_feature
; LOCATION: (319226)..(319226)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (559167)..(559167)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (559241)..(559241)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (600992)..(600992)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (622708)..(622708)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (657081)..(657081)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (657203)..(657203)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (674435)..(674435)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (682442)..(682442)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (713652)..(713652)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (741684)..(741684)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (779455)..(779455)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (779676)..(779676)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (855539)..(855539)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (871619)..(871619)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (1084830)..(1084830)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (1096846)..(1096846)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (1119881)..(1119881)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (1130881)..(1130881)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (1310988)..(1310988)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (1313224)..(1313224)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (1349473)..(1349473)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (1349491)..(1349491)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (1470091)..(1470091)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
```

```
; LOCATION: (1569020)..(1569020)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (1602912)..(1602912)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (1603734)..(1603734)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (1637998)..(1637998)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (1664854)..(1664855)
; OTHER INFORMATION: n equals a, t, c, or g
; US-08-916-421B-1

Query Match 3.5%; Score 37.8; DB 4; Length 1664976;
Best Local Similarity 49.7%; Pred. No. 12;
Matches 96; Conservative 0; Mismatches 97; Indels 0; Gaps 0;

Qy 281 ATTGCTATCTGTACTCCATTGGACTTCTTTACTTACCGGTTTCTACCTATTCCCTGATCT 340
Db 663618 ATTTTTCATCAACCTCATCTGAAATCTGCAAAATACCAAACTACATCATAA 663677
Qy 341 GTGCATCTCAGTTAGCTTCAATGCTTCTTCTTATTTTCTTAACTCACAAAACTTA 400
Db 663678 GAGAGGATATTAGCCCTCATCAATTTCTTCTATATTTTGAATATTCTCTA 663737
Qy 401 CCCCTATCATTTTAAATTTCTTTTCTTCTTAACTATATCTTCCACCTACTTGCATTC 460
Db 663738 AAACCTGGTTTTTAAATTAATTTTCTTCTTCAAGCTCTTATCAACTAACTCAAACTCA 663797
Qy 461 ATAATGAGGAGAC 473
Db 663798 ATATTTAAATAC 663810

RESULT 11
US-09-692-570-1
; Sequence 1, Application US/09692570
; Patent No. 6797466
; GENERAL INFORMATION:
; APPLICANT: Bult et al.
; TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus
; Patent No. 6797466
; TITLE OF INVENTION: jannaschii
; FILE REFERENCE: PB275C1
; CURRENT APPLICATION NUMBER: US/09/692,570
; CURRENT FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: US 60/024,428
; PRIOR FILING DATE: 1996-08-22
; PRIOR APPLICATION NUMBER: US 08/916,421
; PRIOR FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1664976
; TYPE: DNA
; ORGANISM: Methanococcus jannaschii
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (28222)..(28222)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (28257)..(28258)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (84773)..(84773)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
```



```
LOCATION: (84808)..(84808)
FEATURE:
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (84812)..(84812)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (98120)..(98120)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (98159)..(98159)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (98239)..(98239)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (98266)..(98266)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (98343)..(98343)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (103998)..(103998)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (148948)..(148948)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (163385)..(163385)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (191989)..(191989)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (191995)..(191995)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (231980)..(231980)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (234187)..(234187)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (234220)..(234220)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (234814)..(234814)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (309398)..(309398)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (309418)..(309418)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (312837)..(312837)
```

```
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (312993)..(312993)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (319226)..(319226)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (559167)..(559167)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (559241)..(559241)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (600992)..(600992)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (622708)..(622708)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (657081)..(657081)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (657203)..(657203)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (674435)..(674435)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (682442)..(682442)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (713652)..(713652)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (741684)..(741684)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (779455)..(779455)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (779676)..(779676)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (855539)..(855539)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (871619)..(871619)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (1084830)..(1084830)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (1096846)..(1096846)
OTHER INFORMATION: n equals a, t, c, or g
```

```
;
; NAME/KEY: misc_feature
; LOCATION: (1119881)..(1119881)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1110881)..(1130881)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1310988)..(1310988)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1313224)..(1313224)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1349473)..(1349473)
; OTHER INFORMATION: n equals a, t, c, or g

Query Match      3.5%; Score 37.8; DB 4; Length 1664976;
Best Local Similarity 49.7%; Pred. No. 12;
Matches 96; Conservative 0; Mismatches 97; Indels 0; Gaps 0;

Qy 281 ATTGCTATCTGACTCCATTGGACTTCTTTACTTACCCGTTTCTACCTATTCCCTGATCT 340
Db 663618 ATTTTTCATCAACCTCTGTAATCTGCAAAATTACCAATCACCNACTACATCAATA 663677

Qy 341 GTGCATCTCAGTACGCTTCAATGCTTCTTCTCTTATTTTCTTAACCTCACAAAACTTA 400
Db 663678 GAGAGGATATTAAGCCCTCATCAATTTCTTCTATATTTTGTATATTCTCTA 663737

Qy 401 CCCCTATCATTTAAATCTCTTTTCTTCTTAACCTATATCTCCACCCTACTTGCATPCA 460
Db 663738 AAACCTGGTTTTTAATTAATTTTCTTCAAGCTCTTTATCAACTCAACTCAAACTCA 663797

Qy 461 ATAATGAGGAGAC 473
Db 663798 ATATTTAAATAC 663810
```

```
RESULT 12
US-09-007-005-32/c
; Sequence 32, Application US/09007005B
; Patent No. 6258558
; GENERAL INFORMATION:
; APPLICANT: Szostak, Jack W.
; APPLICANT: Roberts, Richard W.
; APPLICANT: Liu, Rihe
; TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN
; FILE REFERENCE: 00786/350003
; CURRENT APPLICATION NUMBER: US/09/007,005B
; EARLIER FILING DATE: 1998-01-14
; EARLIER APPLICATION NUMBER: 60/035,963
; EARLIER FILING DATE: 1997-01-27
; EARLIER APPLICATION NUMBER: 60/064,491
; EARLIER FILING DATE: 1997-11-06
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 32
; LENGTH: 248
; TYPE: RNA
; ORGANISM: Homo sapiens
US-09-007-005-32

Query Match      3.4%; Score 36.8; DB 3; Length 248;
Best Local Similarity 21.0%; Pred. No. 0.2;
Matches 44; Conservative 77; Mismatches 89; Indels 0; Gaps 0;

Qy 237 CCGTGATTGGTTTACGTAGTCTTGACCTTCTTGTAGGACGAGATTGCTATCTGACTC 296
Db 240 CYAVAGYAVGYTGYTAYCYGVCYAVGYCVTYGTYTTCYCYAYGYTGYTGYTGYTGY 181

Qy 297 CATGGACTTCTTACTTACCCGTTTCTACCTATTCCTGATCTGTCATCTCAGTTAGC 356
Db 180 TTYCYAYGYCYTGYTGYTTCYAYCYGVCYAVGYCVTYGTYTTCYCYAYGYTGYTGY 121

Qy 357 CTTCAGATCTTCTCTCTTATTTTCTTAACCTCACAAAACTTACCCCTATCATTTTAAA 416
Db 120 TYCYTTCYCYTTCYCYAYGYAYGYTTCYCYAYGYTTCYCYAYGYTTCYCYTTCYCY 61

Qy 417 TTCTCTTTTCTCTTAATACTATATCTCCAC 446
Db 60 GYCYCYAVTYTGYTAYAVTYTGYTAY 31

RESULT 14
US-09-007-005-3/c
; Sequence 3, Application US/09007005B
; Patent No. 6258558
; GENERAL INFORMATION:
; APPLICANT: Szostak, Jack W.
; APPLICANT: Roberts, Richard W.
; APPLICANT: Liu, Rihe
; TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN
; FILE REFERENCE: 00786/350003
; CURRENT APPLICATION NUMBER: US/09/007,005B
; EARLIER FILING DATE: 1998-01-14
; EARLIER APPLICATION NUMBER: 60/035,963
; EARLIER FILING DATE: 1997-01-27
; EARLIER APPLICATION NUMBER: 60/064,491
; EARLIER FILING DATE: 1997-11-06
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 32
; LENGTH: 248
; TYPE: RNA
; ORGANISM: Homo sapiens
US-09-007-005-32

Query Match      3.4%; Score 36.8; DB 3; Length 248;
Best Local Similarity 21.0%; Pred. No. 0.2;
Matches 44; Conservative 77; Mismatches 89; Indels 0; Gaps 0;

Qy 237 CCGTGATTGGTTTACGTAGTCTTGACCTTCTTGTAGGACGAGATTGCTATCTGACTC 296
Db 240 CYAVAGYAVGYTGYTAYCYGVCYAVGYCVTYGTYTTCYCYAYGYTGYTGYTGYTGY 181

Qy 297 CATGGACTTCTTACTTACCCGTTTCTACCTATTCCTGATCTGTCATCTCAGTTAGC 356
Db 180 TTYCYAYGYCYTGYTGYTTCYAYCYGVCYAVGYCVTYGTYTTCYCYAYGYTGYTGY 121

Qy 357 CTTCAGATCTTCTCTCTTATTTTCTTAACCTCACAAAACTTACCCCTATCATTTTAAA 416
Db 120 TYCYTTCYCYTTCYCYAYGYAYGYTTCYCYAYGYTTCYCYAYGYTTCYCYTTCYCY 61

Qy 417 TTCTCTTTTCTCTTAATACTATATCTCCAC 446
Db 60 GYCYCYAVTYTGYTAYAVTYTGYTAY 31
```



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 1, 2004, 02:40:28 ; Search time 567.648 Seconds  
(without alignments)  
9996.736 Million cell updates/sec

Title: US-09-913-767-6  
Perfect score: 1081  
Sequence: 1 caaatccaacagtccaagat.....ataagctgatatttgaag 1081

Scoring table: IDENTITY NUC  
Gapop 10\_0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues  
Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N Geneseq\_23Sep04:\*  
1: Geneseqn1980s:\*  
2: Geneseqn1990s:\*  
3: Geneseqn2000s:\*  
4: Geneseqn2001as:\*  
5: Geneseqn2001bs:\*  
6: Geneseqn2002as:\*  
7: Geneseqn2002bs:\*  
8: Geneseqn2003as:\*  
9: Geneseqn2003bs:\*  
10: Geneseqn2003cs:\*  
11: Geneseqn2003ds:\*  
12: Geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1077.8	99.7	1081	3	Aaa97924 A. thalia
2	1043.8	96.6	1047	6	Adg88190 A. thalia
3	835.8	77.3	1071	3	Aaa97925 A. thalia
4	406.8	37.6	1293	3	Aaa97922 A. thalia
5	406.2	37.6	3387	6	Abz14407 Arabidops
6	406.2	37.6	3387	8	Ada68246 Arabidops
7	369.6	34.2	1194	3	Aaa97923 A. thalia
8	300	27.8	1295	3	Aac37289 Arabidops
9	298.8	27.6	2660	3	Aac47974 Arabidops
10	298.4	27.6	1145	3	Aaa97921 A. thalia
11	298.4	27.6	1292	3	Aac51119 Arabidops
12	234.2	21.7	83698	6	Abn85767 Arabidops
13	225.8	20.9	1403	3	Aac37652 Arabidops
14	222.8	20.6	2175	8	Ada70748 Rice gene
15	211.6	19.6	1152	8	Ada70213 Rice gene
16	145.6	13.5	1049	3	Aaa97920 A. thalia
17	138.6	12.8	592	10	Abx56692 Arabidops
18	136.6	12.6	2586	3	Aac44184 Arabidops
19	136	12.6	1225	3	Aaa97919 Arabidops
20	129.2	12.0	1155	8	Ada69686 Rice gene
21	127	11.7	1040	8	Ada69568 Rice gene

22	126.4	11.7	332	12	ADP92185	Adp92185 Cotton ex
23	72.2	6.7	774	6	ABN99127	Abn99127 Arabidops
24	66.6	6.2	477	3	AAC37273	Aac37273 Arabidops
25	66.2	6.1	386	6	ABO85215	Abg85215 Arabidops
26	60.8	5.6	1173	12	ADN72520	Adn72520 Thale cre
27	60	5.6	1200	8	ADA70627	Ada70627 Rice gene
28	57.6	5.3	1421	3	AAC39779	Aac39779 Arabidops
29	56	5.2	1418	3	AAC46395	Aac46395 Arabidops
30	51.6	4.8	2000	8	ADA71938	Ada71938 Rice gene
31	50.4	4.7	493	3	AAC36831	Aac36831 Arabidops
32	47.8	4.4	2000	8	ADA71938	Ada71938 Rice gene
33	47.4	4.4	492	3	AAC36845	Aac36845 Arabidops
34	42	3.9	6067	6	ABN80088	Abn80088 Human che
35	42	3.9	110000	6	ABS55320_0	Abss55320 Human tra
36	42	3.9	341511	6	ABS55320	Abss55320 Genomic D
37	40.8	3.8	16633	6	ABN79985	Abn79985 Human che
38	40.6	3.8	72604	2	AAZ10752	Aaz10752 Genomic s
39	40.6	3.8	72604	6	ABK43231	Abk43231 Human HKN
40	40.4	3.7	4590	5	AAH24065	Aah24065 Yeast AOD
41	39.8	3.7	264	12	ACH80963	Ach80963 Human gen
42	39.8	3.7	526	12	ACH67258	Ach67258 Human gen
43	39.8	3.7	5520	5	AAC68924	Aac68924 FLR1 rece
44	39.4	3.6	6467	6	ABN80169	Abn80169 Human che
45	39.2	3.6	446	3	AAC37188	Aac37188 Arabidops

ALIGNMENTS

RESULT 1  
AAA97924  
ID AAA97924 standard; DNA; 1081 BP.  
AC AAA97924;  
XX  
DT 19-JAN-2001 (first entry)  
XX  
DE A. thaliana PUP1 DNA #6.  
XX  
KW PUP1; transgenic plant; nucleobase transporter; apical dominance;  
KW flowering behaviour; senescence; pesticide distribution; ds.  
XX  
OS Arabidopsis thaliana.  
XX  
PN DE19907209-A1.  
XX  
PD 24-AUG-2000.  
XX  
PF 19-FEB-1999; 99DE-01007209.  
XX  
PR 19-FEB-1999; 99DE-01007209.  
XX (FROM/) FROMMER W.  
PI Gillissen B, Buerkle L, Andre B, Frommer WB;  
DR WPI; 2000-566202/53.  
XX  
PT Nucleic acid, useful for producing transgenic plants with altered  
PT nucleobase transporter, encodes a nucleobase transporter protein of  
PT Arabidopsis thaliana.  
XX  
PS Claim 2; Page 14-15; 24pp; German.  
XX  
CC This invention describes a novel nucleic acid encoding a plant nucleobase  
CC transporter (I). (I) is produced by complementation of a nucleobase  
CC transport (NBT)-defective host cell with a plant gene bank by selection  
CC of NBT-positive cells. (I) is used to isolate homologous sequences from  
CC bacteria, fungi, plants, animals and humans, for expression of the  
CC encoded protein (II) in prokaryotic or eukaryotic cells, for inhibiting  
CC expression of (II) (when in antisense orientation), and to produce  
CC transgenic crop plants. The transgenic plants have modified nucleobase  
CC transport properties, e.g. altered affinity and substrate specificity

CC that may result in more efficient nucleobase transport in leaves, changes  
CC in apical dominance, flowering behaviour and senescence, or improved  
CC distribution of pesticides. This sequence encodes the Arabidopsis  
CC thaliana PUP1 protein which is described in the method of the invention  
XX  
SQ Sequence 1081 BP; 322 A; 231 C; 205 G; 323 T; 0 U; 0 Other;

Query Match 99.7%; Score 1077.8; DB 3; Length 1081;  
Best Local Similarity 99.8%; Pred. NO. 8e-297;  
Matches 1079; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy	1	CAAAATCCAAAGTTCAAGATGAAGAAATTCAGTCAGTAGCAGCAAGCAGAAGTATCTC	60
Db	1	CAATCCAAAGTTCAAGATGAAGAAATTCAGTCAGTAGCAGCAAGCAGAAGTATCTC	60
Qy	61	ACTCTAACACATACAAACGGTGGCTCAGGAGTATATATGACAAACGGAGGAAACAGTAA	120
Db	61	ACTCTAACACATACAAACGGTGGCTCAGGAGTATATATGACAAACGGAGGAAACAGTAA	120
Qy	121	TGGCTAGCAACGGTAGTTCAACTTGTGGCTTTCTGTGCTACTTCCATATATATCTTG	180
Db	121	TGGCTAGCAACGGTAGTTCAACTTGTGGCTTTCTGTGCTACTTCCATATATATCTTG	180
Qy	181	TCATTTAAACACATGCAACAACTGATAGATGGAAGAAAGAACTCACCTAGGAACCGT	240
Db	181	TCATTTAAACACATGCAACAACTGATAGATGGAAGAAAGAACTCACCTAGGAACCGT	240
Qy	241	GTATTTGGTTTACGTAGTCTTGGACTTCTTGTAGGAGCAGATTGCTATCTGTACTCCATT	300
Db	241	GTATTTGGTTTACGTAGTCTTGGACTTCTTGTAGGAGCAGATTGCTATCTGTACTCCATT	300
Qy	301	GGACTTCTTTACTTACCCGTTTCTACCTATTCCCTGATCTGTGCATCTCAGTTAGCCCTTC	360
Db	301	GGACTTCTTTACTTACCCGTTTCTACCTATTCCCTGATCTGTGCATCTCAGTTAGCCCTTC	360
Qy	361	AATGCTTTCTTCTCTTATTTCTTAACCTCAACAAACCTTACCCCTATCATTTTAAATCT	420
Db	361	AATGCTTTCTTCTCTTATTTCTTAACCTCAACAAACCTTACCCCTATCATTTTAAATCT	420
Qy	421	CTTTTCTCTTAACTATATCTTCCACCTACTTGCATTCATTAATGAGGACAGACTCC	480
Db	421	CTTTTCTCTTAACTATATCTTCCACCTACTTGCATTCATTAATGAGGACAGACTCC	480
Qy	481	ACAAAGTTACAAAGGAGATGTCAAAGTTTCATATGCACGGTTGCTGCTCTGCT	540
Db	481	ACAAAGTTACAAAGGAGATGTCAAAGTTTCATATGCACGGTTGCTGCTCTGCT	540
Qy	541	GGTTATGCTAGTCTTATCCCTACACAGCTAGCCCTTTCTAAAGTCTCTAAAGACGAA	600
Db	541	GGTTATGCTAGTCTTATCCCTACACAGCTAGCCCTTTCTAAAGTCTCTAAAGACGAA	600
Qy	601	AATTTCTCAGAGTTATCGATATGATATATCTACGTAGTCTAGTGCCAGTTGTTAGC	660
Db	601	AATTTCTCAGAGTTATCGATATGATATCTACGTAGTCTAGTGCCAGTTGTTAGC	660
Qy	661	GTGTGGGGCTTTTGTCTAGCAGTCAGTGGGAAACCTTGACAGTGAATGGATACTAC	720
Db	661	GTGTGGGGCTTTTGTCTAGCAGTCAGTGGGAAACCTTTGACAGTGAATGGATACTAC	720
Qy	721	AAACATGGGAAGGTATCTCATATATGAACCTAGTGTGACAGCTGTTTACCTGGCAGTTA	780
Db	721	AAACATGGGAAGGTATCTCATATATGAACCTAGTGTGACAGCTGTTTACCTGGCAGTTA	780
Qy	781	TTCTCCATCGGTGGCAGGAGTGTATCTTCCAGCTCTCCTCTTATTTCTCAAAATGCAATA	840
Db	781	TTCTCCATCGGTGGCAGGAGTGTATCTTCCAGCTCTCCTCTTATTTCTCAAAATGCAATA	840
Qy	841	AGCGTTTGGGACTCCAGTGGTTCCTATCTTGGCTGTAAATCATTTTCCATGACAAATG	900
Db	841	AGCGTTTGGGACTCCAGTGGTTCCTATCTTGGCTGTAAATCATTTTCCATGACAAATG	900
Qy	901	AATGGTTTAAAGGTGATTTCTATGATCTTACCTAGTATTTGGGGTTTCACCTTCTATGCTAC	960
Db	901	AATGGTTTAAAGGTGATTTCTATGATCTTACCTAGTATTTGGGGTTTCACCTTCTATGCTAC	960

Db	901	AATGGTTTAAAGGTGATTTCTATGATCTTACCTAGTATTTGGGGTTTCACCTTCTATGCTAC	960
Qy	961	CACAAATATCTTGTATGACAAAACTTTGAGAAAAATCATGAAATCACAACACAGATCC	1020
Db	961	CAACAATATCTTGTATGACAAAACTTTGAGAAAAATCATGAAATCACAACACAGATCC	1020
Qy	1021	CCTGACCCACCAAGCAGAGAGTCAACTTTGGCAATCAAAAATAAGCTGATATTTTCAAA	1080
Db	1021	CCTGACCCACCAAGCAGAGAGTCAACTTTGGCAATCAAAAATAAGCTGATATTTTCAAA	1080
Qy	1081	G 1081	
Db	1081	G 1081	
RESULT 2			
ADG88190			
ID	ADG88190	standard; cDNA; 1047 BP.	
XX	AC	ADG88190;	
XX	DT		
XX	22-APR-2004	(first entry)	
XX	DE	A. thaliana RPP4-upregulated pathogen infection-related gene #632.	
XX	KW	Pathogen infection-related gene; Plant; Peronospora parasitica;	
KW	KW	defence mechanism; RPP4; pathogen resistance; transgenic plant; oomycete;	
KW	KW	fungus; bacterium; virus; nematode; insect; aphid; gene; ss.	
XX	OS	Arabidopsis thaliana.	
XX	PN	WO200222675-A2.	
XX	PD	21-MAR-2002.	
XX	14-SEP-2001;	2001WO-US028506.	
XX	15-SEP-2000;	2000US-0232778P.	
PR	22-JUN-2001;	2001US-0300183P.	
XX	(SYGN )	SYNGENTA PARTICIPATIONS AG.	
PA	(UNYC-)	UNIV NORTH CAROLINA.	
PA	(GLAZ/)	GLAZEBROOK J.	
PA	(WANG/)	WANG X.	
PA	(DANG/)	DANGL J L.	
PA	(EULG/)	EULGEM T.	
PA	(ZHUT/)	ZHU T.	
XX	Glazebrook J,	Wang X, Dangl JL, Eulgem T, Zhu T;	
PI	WPI;	2002-292409/33.	
XX	Novel isolated polynucleotide,	useful for conveying pathogen resistance	
XX	to plants,	and for identifying plants infected with a pathogen.	
PT	Claim 3;	SEQ ID NO 632; 605pp; English.	
PS	The invention relates to 691 Arabidopsis thaliana genes (ADG87559--		
CC	ADG87557))	whose expression is altered in response to pathogen infection,	
CC	and to homologues of these genes	from other plants or fungi, especially	
CC	from maize, soybean, barley, alfalfa, sunflower, canola (oilseed rape),		
CC	cotton, peanut, sorghum, tobacco, sugarbeet, rice or wheat. The		
CC	expression of genes of the invention was upregulated or downregulated in		
CC	Arabidopsis plants infected with the oomycete Peronospora parasitica,		
CC	indicating that they play a role in defence mechanisms. The genes of the		
CC	invention are regulated by RPP7 or RPP8 which act via unconventional		
CC	signalling cascades, or by the RPP4-dependent pathway. The invention also		
CC	relates to polypeptides encoded by the pathogen infection-related genes;		
CC	promoter motifs from pathogen infection-related genes (ADG88243-ADG88327)		
CC	; expression cassettes, host cells and pathogen-resistant transgenic		
CC	plants and their progeny comprising a polynucleotide of the invention;		
CC	and a method of identifying a plant cell infected with a pathogen. The		
CC	polynucleotide sequences and methods of the invention are useful for		

CC identifying plants infected with a pathogen, and for conferring  
CC resistance to pathogens such as oomycetes, fungi, bacteria, viruses,  
CC nematodes and insects (e.g., aphids). The present sequence represents an  
CC Arabidopsis thaliana gene whose expression is altered in response to  
CC Peronospora parasitica infection. Note: The sequence data for this patent  
CC did not form part of the printed specification, but was obtained in  
CC electronic format directly from wipo at  
CC ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 1047 BP; 309 A; 225 C; 199 G; 314 T; 0 U; 0 Other;

Query Match 96.6%; Score 1043.8; DB 6; Length 1047;

Best Local Similarity 99.8%; Pred. No. 4e-287;

Matches 1045; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 19 ATGAAAGAAATTCAGTCAGTAGCAGCAAGCAGCAAGTATCTCACTTAACACATACAAAC 78  
DB 1 ATGAAAGAAATTCAGTCAGTAGCAGCAAGCAGCAAGTATCTCACTTAACACATACAAAC 60  
QY 79 GGTGGCTCAGGAGTATATATATGACAAACGAGGAAACAGTAAATGGCTAGCAACGGTAGTT 138  
DB 61 GGTGGCTCAGGAGTATATATGACAAACGAGGAAACAGTAAATGGCTAGCAACGGTAGTT 120  
QY 139 CAACTTGTGGCTTCTCTGTGCTACTTCCATATATATATCTGTCTATTTAAACACATGCA 198  
DB 121 CAACTTGTGGCTTCTCTGTGCTACTTCCATATATATATCTGTCTATTTAAACACATGCA 180  
QY 199 ACAACTGATAGATGGAAGAAAGAACTCACCTAGGAAACCGTGTATTGGTTTACGTTAG 258  
DB 181 ACACTGATAGATGGAAGAAAGAACTCACCTAGGAAACCGTGTATTGGTTTACGTTAG 240  
QY 259 CTTGGACTTCTTGTAGGAGCAGATTCTATCTGTACTCCATTGGACTCTTTTACTTACCC 318  
DB 241 CTTGGACTTCTTGTAGGAGCAGATTCTATCTGTACTCCATTGGACTCTTTTACTTACCC 300  
QY 319 GTTCTTACCTATTCCTGATCTGTGCACTCTAGTTAGCCTTCAATGCTTTCTTCTTAT 378  
DB 301 GTTCTTACCTATTCCTGATCTGTGCACTCTAGTTAGCCTTCAATGCTTTCTTCTTAT 360  
QY 379 TTCTTTAACTCAAAAACCTTACCCTATCATTTTAAATTTCTTTTCTCTTAACTATA 438  
DB 361 TTCTTTAACTCAAAAACCTTACCCTATCATTTTAAATTTCTTTTCTCTTAACTATA 420  
QY 439 TCTTCCACCTACTTGCATTCATATATGAGGACAGACTCCAAAAGTTTCAAAAGA 498  
DB 421 TCTTCCACCTACTTGCATTCATATATGAGGACAGACTCCAAAAGTTTCAAAAGA 480  
QY 499 GAGTATGTCAAAGTTTCATATGCACCGTTGCTGCTGCTGCTGCTGCTGCTGCTGCT 558  
DB 481 GAGTATGTCAAAGTTTCATATGCACCGTTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540  
QY 559 TCCCTACACAGCTAGCCCTTTCTTAAAGTCTTAAAGTCTTAAAGTCTTAAAGTCTTAA 618  
DB 541 TCCCTACACAGCTAGCCCTTTCTTAAAGTCTTAAAGTCTTAAAGTCTTAAAGTCTTAA 600  
QY 619 GATATGATAATCTACGTGAGTCTAGTGGCAGTTGTTAGGCTGGTGGGCTTTTGGT 678  
DB 601 GATATGATAATCTACGTGAGTCTAGTGGCAGTTGTTAGGCTGGTGGGCTTTTGGT 660  
QY 679 AGCAGTGTGGAATACTTTGAGCAGTGAATGGATACTACAAACATGGGAAGGTATCC 738  
DB 661 AGCAGTGTGGAATACTTTGAGCAGTGAATGGATACTACAAACATGGGAAGGTATCC 720  
QY 739 TACATTATCAACTAGTGGGACAGCTGTTACTGCGAGTTATCTCCATCGGTGGCACA 798  
DB 721 TACATTATCAACTAGTGGGACAGCTGTTACTGCGAGTTATCTCCATCGGTGGCACA 780  
QY 799 GGACTGATCTTGGAGCTCTCTCTATTTCTCAATGCAATAGCGTTTGGGACTCCCA 858  
DB 781 GGACTGATCTTGGAGCTCTCTCTATTTCTCAATGCAATAGCGTTTGGGACTCCCA 840  
QY 859 GTGGTTCCTATCTTGGCTGTAACTATTTTCCATGACAAATGAATGGGTTAAAGGTGATT 918  
DB 841 GTGGTTCCTATCTTGGCTGTAACTATTTTCCATGACAAATGAATGGGTTAAAGGTGATT 900  
QY 919 TCTATGATCCTAGCTATTTGGGGTTTCACTTCTTGTCTACCAACAATATCTTGTATGAC 978  
DB 901 TCTATGATCCTAGCTATTTGGGGTTTCACTTCTTGTCTACCAACAATATCTTGTATGAC 960  
QY 979 AAAAACTTGAAGAAAAAATCATGAATCAACAACAGAAATCCCTGACCCACCAAGCA 1038  
DB 961 AAAAACTTGAAGAAAAAATCATGAATCAACAACAGAAATCCCTGACCCACCAAGCA 1020  
QY 1039 GAAGAGTCAACTTGGCAATCAAAATAA 1065  
DB 1021 GAAGAGTCAACTTGGCAATCAAAATAA 1047

## RESULT 3

AAA97925

ID AAA97925 standard; DNA; 1071 BP.

XX AC AAA97925;

XX DT 19-JAN-2001 (first entry)

XX DE A. thaliana PUP1 DNA #7.

XX KW PUP1; transgenic plant; nucleobase transporter; apical dominance;

XX KW flowering behaviour; senescence; pesticide distribution; ds.

XX OS Arabidopsis thaliana.

XX PN DE19907209-A1.

XX PD 24-AUG-2000.

XX PF 19-FEB-1999; 99DE-01007209.

XX PR 19-FEB-1999; 99DE-01007209.

XX PA (FROM/) FROMMER W.

XX PI Gillissen B, Buerkle L, Andre B, Frommer WB;

XX WPI; 2000-566202/53.

XX Nucleic acid, useful for producing transgenic plants with altered

XX nucleobase transport, encodes a nucleobase transporter protein of

XX Arabidopsis thaliana.

XX Claim 2; Page 15; 24pp; German.

CC This invention describes a novel nucleic acid encoding a plant nucleobase  
CC transporter (I). (I) is produced by complementation of a nucleobase  
CC transport (NBT)-defective host cell with a plant gene bank by selection  
CC of NBT-positive cells. (I) is used to isolate homologous sequences from  
CC bacteria, fungi, plants, animals and humans, for expression of the  
CC encoded protein (II) in prokaryotic or eukaryotic cells, for inhibiting  
CC expression of (II) (when in antisense orientation), and to produce  
CC transgenic crop plants. The transgenic plants have modified nucleobase  
CC transport properties, e.g. altered affinity and substrate specificity  
CC that may result in more efficient nucleobase transport in leaves, changes  
CC in apical dominance, flowering behaviour and senescence, or improved  
CC distribution of pesticides. This sequence encodes the Arabidopsis  
CC thaliana PUP1 protein which is described in the method of the invention  
CC  
SQ Sequence 1071 BP; 307 A; 236 C; 208 G; 320 T; 0 U; 0 Other;

Query Match 77.3%; Score 835.8; DB 3; Length 1071;

Best Local Similarity 90.2%; Pred. No. 8.9e-228;

Matches 894; Conservative 0; Mismatches 97; Indels 0; Gaps 0;

QY 89 GAGTATATCTATGACAAACGAGGAAACAGTAAATGGCTAGCAACGGTAGTTCAACTTTG 148

DB 80 GACTATATCTATGAAATGGAGAAACAGCAAAATGGCTAGCAACGGTAGTTCACTTTG 139









DT 19-JAN-2001 (first entry)  
 XX DE A. thaliana PUP1 DNA #5.  
 XX PUP1; transgenic plant; nucleobase transporter; apical dominance;  
 KW flowering behaviour; senescence; pesticide distribution; ds.  
 KW Arabidopsis thaliana.  
 OS  
 XX DE1907209-A1.  
 XX 24-AUG-2000.  
 XX 19-FEB-1999; 99DE-01007209.  
 XX 19-FEB-1999; 99DE-01007209.  
 XX (FROM/) FROMMER W.  
 XX Gillissen B, Buerkle L, Andre B, Frommer WB;  
 PI WPI; 2000-566202/53.  
 XX Nucleic acid, useful for producing transgenic plants with altered  
 PT nucleobase transport, encodes a nucleobase transporter protein of  
 PT Arabidopsis thaliana.  
 XX Claim 1f; Page 14; 24pp; German.  
 XX This invention describes a novel nucleic acid encoding a plant nucleobase  
 CC transporter (I). (I) is produced by complementation of a nucleobase  
 CC of NBT-positive cells. (I) is used to isolate homologous sequences from  
 CC bacteria, fungi, plants, animals and humans, for expression of the  
 CC encoded protein (II) in prokaryotic or eukaryotic cells, for inhibiting  
 CC expression of (II) (when in antisense orientation), and to produce  
 CC transgenic crop plants. The transgenic plants have modified nucleobase  
 CC transport properties, e.g. altered affinity and substrate specificity  
 CC that may result in more efficient nucleobase transport in leaves, changes  
 CC in apical dominance, flowering behaviour and senescence, or improved  
 CC distribution of pesticides. This sequence encodes the Arabidopsis  
 CC thaliana PUP1 protein which is described in the method of the invention  
 XX Sequence 1194 BP; 298 A; 279 C; 231 G; 386 T; 0 U; 0 Other;  
 SQ

Query Match 34.2%; Score 369.6; DB 3; Length 1194;  
 Best Local Similarity 62.2%; Pred. No. 9.9e-95;  
 Matches 601; Conservative 0; Mismatches 359; Indels 6; Gaps 1;  
 1; 197 GATTGTACTATGAAATGAGGAAATAGCATAATGTGTAACACTTCTTCAACTCATTTG 256  
 149 GCTTCTCTGTCTACTTCCATATATATATCTTGTCAATTA-----AAACACATGCAACAA 202  
 257 GCTTCCCTCTACTGGTGTCTGTCGCGCTTCTTCTCGAATCAGGCAACCAATCAACAG 316  
 203 CTGATAGAGATGMAAAGAACTCACCTAGGACCGTGTATTGGTTTACGTAGTCTTG 262  
 317 ATACAAATTCAGTCAGTCCCTTCTTCCACACCCCTTGCATCGGTTTACTTGTGCATG 376  
 263 GACTTCTTCTAGGAGCAGATGCTATCTGTACTTCCATTTGGACTTCTTACTTACCCGTTT 322  
 377 GACTGTAGTGTCCGCTTATGCTTATTTCTGTGCACTAGGTTGCTCTACTTACCACTCT 436  
 323 CTACCTATTTCCCTGATCTGTGATCTCAGTTAGCCTTCAATGCTTCTTCTCTATTTTC 382  
 437 CTACTTTCTCCCTCATCTTGGGCTTCAAGTTGGCCTTCACTGCCTTTTCTCATATTTCC 496  
 383 TTAACCTCACAAAACCTTACCCCTATCATTTTAAATTCCTTTTCTCTTAACTATATCTT 442  
 497 TTAACCTCGGAAAGTTCACTCTTTGATAGTCAGTTCTTTGCTTCTCTCTCACTGATCTCT 556

QY 443 CCACCCCTACTTGCATTCATTAATAGAGAGACAGACTCCACAAAAGTTACAAAAGGAGAGT 502  
 DB 557 CTGCTCTTCTTGTGGTCAACACTGATTCAGAAAACCTCAATATGATCTAGAGTACAGT 616  
 QY 503 ATGTCAAGGTTTCATATGCACCGTGTGCTGCTGCTGTTATGTTCTAGTCTTATCCC 562  
 DB 617 ATGTGATCGGTTTCATATGATACCATCGGTGCTTCCGCTGGGATTTGACTGTACTATCTC 676  
 QY 563 TACAACAGCTAGCTTCTTAAAGTCTTAAAGAGCAAAATTTCTCAGAAAGTTATCGATA 622  
 DB 677 TGATACAAATGCTCTTTCAGGAAAGTTTTCAGAGCATATACCTCAGCAGTACCGACT 736  
 QY 623 TGATAATCTACGTAGTCTAGTGGCCAGTTGTGTAGCGTGGTGGGCTTTTGTCTAGCA 682  
 DB 737 TGGCCATTTACCAGTCTCTAGTTGCGAGTTGTGTAGTTCTCATAGGACTTTTTTGCAGTG 796  
 QY 683 GTGAGTGGAAACTTTTGAGCAGTGAATGGATAACTACAAAACATGGGAAGGTATCTTACA 742  
 DB 797 GAGAGTGGGAAACTTTTGCCAACTGAGATGAGAAACTTACAAACTCGGGAAGGTGTCTATG 856  
 QY 743 TTATGAACCTAGTGTGGACAGCTGTTTACTCTGGCAGTTATTTCTCCATCGGTGGCACAGGAC 802  
 DB 857 TTTTGACTTTAGCTCTGGCAGTATTTCTGGCAAGTCTACACTCTTGGTCTTGTGGGAT 916  
 QY 803 TGATCTTCGAGCTCTCTCTCTTCTTATTTCTCAAATGCAATAGCGTTTTTGGGACTCCCAAGTG 862  
 DB 917 TGATCTTCGAGTCACTCTCTGTTCTTCAATTTCCATAACAGCTGTGGGATTTGCCTATAG 976  
 QY 863 TTCCCTATCTTGGCTGTATCATTTTTTCCATGACAAAATGAATGGTGAAGGTGATTCTTA 922  
 DB 977 TTTCCAGTTTGGCGCAGTGTAGTTTTTCCATGATAGAGACGATCAAAAATCTTCTCCA 1036  
 QY 923 TGATCTCTAGTATTTTGGGTTTTCCTCTATGCTTACCACAAATATCTTGTATGACAAA 982  
 DB 1037 TTTATTTTGTATCTTGGGCTTCTTTCATTCGTCTATCAGACCTACCTTCGACGAAAAGA 1096  
 QY 983 ACTTCAAGAAAATCATGAAATCAACAACAGAAATCCCTGACCCACAGAGCAGAG 1042  
 DB 1097 AGTTGAATACTAGCCACACAGTGTGTAGGAGATCTTCTATCTCTGTTGAGGAAGTTC 1156  
 QY 1043 AGTCAA 1048  
 DB 1157 ACACAA 1162

RESULT 8  
 AAC37289  
 ID AAC37289 standard; DNA; 1295 BP.  
 XX AAC37289;  
 AC AC  
 XX AC  
 DT 17-OCT-2000 (first entry)  
 XX  
 DE Arabidopsis thaliana DNA fragment SEQ ID NO: 16846.  
 XX  
 KW Hybridisation assay; genetic mapping; gene expression control;  
 KW protein identification; signal transduction pathway; metabolic pathway;  
 KW promoter; termination sequence; ss.  
 XX Arabidopsis thaliana.  
 OS  
 XX  
 PN EP1033405-A2.  
 PD 06-SEP-2000.  
 XX  
 PF 25-FEB-2000; 2000EP-00301439.  
 PR 25-FEB-1999; 99US-0121825P.  
 PR 05-MAR-1999; 99US-0123180P.  
 PR 09-MAR-1999; 99US-0123548P.  
 PR 23-MAR-1999; 99US-0125788P.  
 PR 25-MAR-1999; 99US-0126264P.  
 PR 29-MAR-1999; 99US-0126785P.

PR 01-APR-1999; 99US-0137462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-01351124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135639P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-01445086P.  
PR 21-JUL-1999; 99US-01445088P.  
PR 22-JUL-1999; 99US-01445085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 06-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147453P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 25-AUG-1999; 99US-0149930P.  
PR 26-AUG-1999; 99US-0150566P.  
PR 27-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 30-AUG-1999; 99US-0151080P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.

PR	13-OCT-1999;	99US-0159293P.	
PR	13-OCT-1999;	99US-0159294P.	
PR	13-OCT-1999;	99US-0159295P.	
PR	14-OCT-1999;	99US-0159329P.	
PR	14-OCT-1999;	99US-0159330P.	
PR	14-OCT-1999;	99US-0159331P.	
PR	14-OCT-1999;	99US-0159337P.	
PR	14-OCT-1999;	99US-0159638P.	
PR	18-OCT-1999;	99US-0159584P.	
PR	21-OCT-1999;	99US-0160741P.	
PR	21-OCT-1999;	99US-0160767P.	
PR	21-OCT-1999;	99US-0160768P.	
PR	21-OCT-1999;	99US-0160770P.	
PR	21-OCT-1999;	99US-0160814P.	
PR	21-OCT-1999;	99US-0160815P.	
PR	22-OCT-1999;	99US-0160980P.	
PR	22-OCT-1999;	99US-0160981P.	
PR	22-OCT-1999;	99US-0160989P.	
PR	25-OCT-1999;	99US-0161404P.	
PR	25-OCT-1999;	99US-0161405P.	
PR	25-OCT-1999;	99US-0161406P.	
PR	26-OCT-1999;	99US-0161359P.	
PR	26-OCT-1999;	99US-0161360P.	
PR	26-OCT-1999;	99US-0161361P.	
PR	28-OCT-1999;	99US-0161920P.	
PR	28-OCT-1999;	99US-0161992P.	
PR	28-OCT-1999;	99US-0161993P.	
PR	29-OCT-1999;	99US-0162142P.	
Query Match 27.8%; Score 300; DB 3; Length 1295;			
Best Local Similarity 58.4%; Pred. No. 7.4e-75;			
Matches 572; Conservative 0; Mismatches 390; Indels 18; Gaps 2;			
QY	89	GAGTATACATGACACGGAGGAAACAGTAAATGGCTAGCAACGGTAGTTCAACTGTG 148	
DB	192	GACTTTACTAGAAAAGCGGTAAAGCAGCATGGCTCGAACCTTGGTTCAGCTTGTAG 251	
QY	149	GCTTTCCTGTGCTATCCATATATATCTGTGCAITTTAAACACATGCAACAACTGATA 208	
DB	252	GGTTTCCTTTAAACCTTCTCTGTATATTTACTTAAAGCGTGAGCGCTCCAGACTATAA 311	
QY	209	GAGATGGAAGAAAGAACTCACTAG-----GAAACGTGTATGGTTTACGTAGTCTTG 262	
DB	312	CCATTACCAAAAATACTTCTCTCTGACACTATCTTTAGTGTATATTTGGACTTG 371	
QY	263	GACTTCTGTAGGAGCAGATTGCTATCTGTACTCCATTTGACTTCTTTACTTACCGTTT 322	
DB	372	GCTTGTCTGTGTCGACATGTATTTGTACTCAITTTGGCTACTTTACCTTCTGTCT 431	
QY	323	CTACCTATTCCCTGATCTGTGCACTCTCAGTTAGCCTTCAATGCTTTCTTCTCTATTTTC 382	
DB	432	CAACTTCTCTTTGATCTCTGGTGCATTTGGCTTTTAAAGCGCTTCTCTTACTTCTC 491	
QY	383	TAACTCAGAAAACCTACCCCTATCATTTTAAATCTCTTTTCTCTTAACTATATCTT 442	
DB	492	TAACTCAGAAAATAACACACATTTATATCTCAATTCACCTTGTCTCTTAAACATATCT 551	
QY	443	CCACCTCTCTGATTCATTAATGAGGAGACGACCTCCAAAGTTTACAAAGGAG--- 499	
DB	552	CTACCTCTCTGTTATCCACATGAACCGAATCTCCCTCTCTTACTTCAAAAGTCGCAG 611	
QY	500	-----AGTATCAAGAGTTTCATATGCAACCGTTCGTCTGCTGTGTTATGTTTC 550	
DB	612	CCAAGTCCAAGTATGTGATTGGATACATCTGCGCGTTCGTAGCTCAGCTGTATTCTC 671	
QY	551	TAGTCTTATCCCTACAAAGCTAGTCCCTTTCTTAAAGTCTTAAAGAACAAATTTCTCAG 610	
DB	672	TGCTGCTTTCTTTAAACAGATTTAGCGCTTCGAAAAGATTCTTAAAGAAATACATCAAG 731	
QY	611	AAGTTATGATATGATAATCTACGTGAGTCTAGTGGCCAGTTGTGTAGCGTGTGGGC 670	
DB	732	CTATTTTAGACATGCGCCACATATCCGCTCTATGCTAGTACTTGTGTAGTGTGGTAGGAC 791	

QY	671	TTTTTGCTAGCAGTGAGTGGAAAAAAGCTTTGAGCAGTGAAATGGATAACTACAACATGGGA 730	
DB	792	TTTTTGGAAAGTGGTGGTGGAAAAAGCTGAGTACAGAAATGGAAGAGTTTCAACTAGGGA 851	
QY	731	AGGTATCCTACATATTAGAACCTAGTGTGACAGCTGTTACCTGGCAGTTATTCTCCATCG 790	
DB	852	AAAGCTCATACATTTTGATAAACAATCGGTTCAACGATATCATGGCAAGCTTGTGTTGATTG 911	
QY	791	GTGGCACAGGACTGATCTTCGAGCTCTCTCTCTATTTCTCAAAATGCAATGAAGCGTTTGG 850	
DB	912	GAAGTGTGTTGATTATCGAAGTTTCAATCGCTTTTCCATGTGTCATAGCACTCTTT 971	
QY	851	GACTCCAGTGGTTCCTATCTTTGGCTGTAATCAATTTCCATGACAAAAATGAATGGGTAA 910	
DB	972	GTTTACCAGTGTGCTGTTCTTCTGCTGTTCTTCTTCGATGAGATGAGTGGGAATCA 1031	
QY	911	AGGTGATTTCTATGATCTTAGCTATTTGGGGTTTCACTTCTCTATGTCTTACCAATATC 970	
DB	1032	AGTTGATTGCAATGTTTTTGGCCATCTGGGGATTTGTTTCTTATGTTATCAGCATATG 1091	
QY	971	TTGATGACAAAAAATCTGAAGAAAAATCATGAAATCACAACAAGAAATCCCTGACCCAC 1030	
DB	1092	TCAATGATAGAAAGCCAGAGAGAGACCAAGAGCTTCTCAGTCTAAAGAGAGAAGAAC 1151	
QY	1031	CAGAAGCAGAGAGTCAACT 1050	
DB	1152	AAAAACAAGTAGATACCAT 1171	
RESULT 9			
AAC47974			
ID	AAC47974 standard; DNA; 2660 BP.		
XX	AAC47974;		
XX	18-OCT-2000 (first entry)		
DE	Arabidopsis thaliana DNA fragment SEQ ID NO: 55799.		
KW	Hybridisation assay; genetic mapping; gene expression control;		
KW	protein identification; signal transduction pathway; metabolic pathway;		
XX	promoter; termination sequence; ss.		
OS	Arabidopsis thaliana.		
PN	EP1033405-A2.		
PD	06-SEP-2000.		
XX	25-FEB-2000; 2000BP-00301439.		
XX	25-FEB-1999; 99US-0121825P.		
PR	05-MAR-1999; 99US-0123180P.		
PR	09-MAR-1999; 99US-0123548P.		
PR	23-MAR-1999; 99US-0125788P.		
PR	25-MAR-1999; 99US-0126264P.		
PR	29-MAR-1999; 99US-0126785P.		
PR	01-APR-1999; 99US-0127462P.		
PR	06-APR-1999; 99US-0128234P.		
PR	08-APR-1999; 99US-0128714P.		
PR	16-APR-1999; 99US-0129845P.		
PR	19-APR-1999; 99US-0130077P.		
PR	21-APR-1999; 99US-0130449P.		
PR	23-APR-1999; 99US-0130510P.		
PR	23-APR-1999; 99US-0130891P.		
PR	28-APR-1999; 99US-0131449P.		
PR	30-APR-1999; 99US-0132048P.		
PR	30-APR-1999; 99US-0132407P.		
PR	04-MAY-1999; 99US-0132484P.		
PR	05-MAY-1999; 99US-0132485P.		
PR	06-MAY-1999; 99US-0132486P.		
PR	06-MAY-1999; 99US-0132487P.		
PR	07-MAY-1999; 99US-0132863P.		

PR	11-MAY-1999;	99US-0134256P.
PR	14-MAY-1999;	99US-0134218P.
PR	14-MAY-1999;	99US-0134219P.
PR	14-MAY-1999;	99US-0134221P.
PR	14-MAY-1999;	99US-0134370P.
PR	18-MAY-1999;	99US-0134768P.
PR	19-MAY-1999;	99US-0134941P.
PR	20-MAY-1999;	99US-0135124P.
PR	21-MAY-1999;	99US-0135353P.
PR	24-MAY-1999;	99US-0135629P.
PR	25-MAY-1999;	99US-0136021P.
PR	27-MAY-1999;	99US-0136392P.
PR	28-MAY-1999;	99US-0136782P.
PR	01-JUN-1999;	99US-0137222P.
PR	03-JUN-1999;	99US-0137528P.
PR	04-JUN-1999;	99US-0137502P.
PR	07-JUN-1999;	99US-0137724P.
PR	08-JUN-1999;	99US-0138094P.
PR	10-JUN-1999;	99US-0138540P.
PR	10-JUN-1999;	99US-0138847P.
PR	14-JUN-1999;	99US-0139119P.
PR	16-JUN-1999;	99US-0139452P.
PR	16-JUN-1999;	99US-0139453P.
PR	17-JUN-1999;	99US-0139492P.
PR	18-JUN-1999;	99US-0139454P.
PR	18-JUN-1999;	99US-0139455P.
PR	18-JUN-1999;	99US-0139456P.
PR	18-JUN-1999;	99US-0139457P.
PR	18-JUN-1999;	99US-0139458P.
PR	18-JUN-1999;	99US-0139459P.
PR	18-JUN-1999;	99US-0139460P.
PR	18-JUN-1999;	99US-0139461P.
PR	18-JUN-1999;	99US-0139462P.
PR	18-JUN-1999;	99US-0139463P.
PR	18-JUN-1999;	99US-0139750P.
PR	18-JUN-1999;	99US-0139763P.
PR	21-JUN-1999;	99US-0139817P.
PR	22-JUN-1999;	99US-0139899P.
PR	23-JUN-1999;	99US-0140353P.
PR	23-JUN-1999;	99US-0140354P.
PR	24-JUN-1999;	99US-0140695P.
PR	28-JUN-1999;	99US-0140823P.
PR	29-JUN-1999;	99US-0140991P.
PR	30-JUN-1999;	99US-0141287P.
PR	01-JUL-1999;	99US-0141842P.
PR	01-JUL-1999;	99US-0142154P.
PR	02-JUL-1999;	99US-0142055P.
PR	06-JUL-1999;	99US-0142302P.
PR	08-JUL-1999;	99US-0142803P.
PR	09-JUL-1999;	99US-0142920P.
PR	12-JUL-1999;	99US-0142977P.
PR	13-JUL-1999;	99US-0143542P.
PR	14-JUL-1999;	99US-0143624P.
PR	15-JUL-1999;	99US-0144005P.
PR	16-JUL-1999;	99US-0144085P.
PR	16-JUL-1999;	99US-0144086P.
PR	19-JUL-1999;	99US-0144325P.
PR	19-JUL-1999;	99US-0144331P.
PR	19-JUL-1999;	99US-0144332P.
PR	19-JUL-1999;	99US-0144333P.
PR	19-JUL-1999;	99US-0144334P.
PR	19-JUL-1999;	99US-0144335P.
PR	20-JUL-1999;	99US-0144352P.
PR	20-JUL-1999;	99US-0144632P.
PR	20-JUL-1999;	99US-0144884P.
PR	21-JUL-1999;	99US-0144814P.
PR	21-JUL-1999;	99US-0145086P.
PR	21-JUL-1999;	99US-0145088P.
PR	22-JUL-1999;	99US-0145085P.
PR	22-JUL-1999;	99US-0145087P.
PR	22-JUL-1999;	99US-0145089P.
PR	22-JUL-1999;	99US-0145192P.
PR	23-JUL-1999;	99US-0145145P.
PR	23-JUL-1999;	99US-0145145P.
PR	23-JUL-1999;	99US-0145218P.
PR	23-JUL-1999;	99US-0145224P.
PR	26-JUL-1999;	99US-0145276P.
PR	27-JUL-1999;	99US-0145913P.
PR	27-JUL-1999;	99US-0145918P.
PR	27-JUL-1999;	99US-0145919P.
PR	28-JUL-1999;	99US-0145951P.
PR	02-AUG-1999;	99US-0146386P.
PR	02-AUG-1999;	99US-0146388P.
PR	02-AUG-1999;	99US-0146389P.
PR	02-AUG-1999;	99US-0146389P.
PR	03-AUG-1999;	99US-0147038P.
PR	04-AUG-1999;	99US-0147204P.
PR	04-AUG-1999;	99US-0147302P.
PR	05-AUG-1999;	99US-0147192P.
PR	05-AUG-1999;	99US-0147260P.
PR	06-AUG-1999;	99US-0147303P.
PR	06-AUG-1999;	99US-0147416P.
PR	09-AUG-1999;	99US-0147493P.
PR	09-AUG-1999;	99US-0147935P.
PR	10-AUG-1999;	99US-0148171P.
PR	11-AUG-1999;	99US-0148319P.
PR	12-AUG-1999;	99US-0148341P.
PR	13-AUG-1999;	99US-0148565P.
PR	13-AUG-1999;	99US-0148684P.
PR	16-AUG-1999;	99US-0149368P.
PR	17-AUG-1999;	99US-0149175P.
PR	18-AUG-1999;	99US-0149426P.
PR	20-AUG-1999;	99US-0149722P.
PR	20-AUG-1999;	99US-0149723P.
PR	20-AUG-1999;	99US-0149929P.
PR	23-AUG-1999;	99US-0149902P.
PR	23-AUG-1999;	99US-0149930P.
PR	25-AUG-1999;	99US-0150566P.
PR	26-AUG-1999;	99US-0150884P.
PR	27-AUG-1999;	99US-0151065P.
PR	27-AUG-1999;	99US-0151066P.
PR	30-AUG-1999;	99US-0151303P.
PR	31-AUG-1999;	99US-0151438P.
PR	01-SEP-1999;	99US-0151930P.
PR	07-SEP-1999;	99US-0152363P.
PR	10-SEP-1999;	99US-0153070P.
PR	13-SEP-1999;	99US-0153758P.
PR	15-SEP-1999;	99US-0154018P.
PR	16-SEP-1999;	99US-0154039P.
PR	20-SEP-1999;	99US-0154779P.
PR	22-SEP-1999;	99US-0155139P.
PR	23-SEP-1999;	99US-0155486P.
PR	24-SEP-1999;	99US-0155659P.
PR	28-SEP-1999;	99US-0156458P.
PR	29-SEP-1999;	99US-0156596P.
PR	04-OCT-1999;	99US-0157117P.
PR	05-OCT-1999;	99US-0157753P.
PR	06-OCT-1999;	99US-0157865P.
PR	07-OCT-1999;	99US-0158029P.
PR	08-OCT-1999;	99US-0158232P.
PR	12-OCT-1999;	99US-0158369P.
PR	13-OCT-1999;	99US-0159293P.
PR	13-OCT-1999;	99US-0159294P.
PR	13-OCT-1999;	99US-0159295P.
PR	14-OCT-1999;	99US-0159329P.
PR	14-OCT-1999;	99US-0159330P.
PR	14-OCT-1999;	99US-0159331P.
PR	14-OCT-1999;	99US-0159637P.
PR	14-OCT-1999;	99US-0159638P.
PR	18-OCT-1999;	99US-0159584P.
PR	21-OCT-1999;	99US-0160741P.
PR	21-OCT-1999;	99US-0160767P.
PR	21-OCT-1999;	99US-0160768P.
PR	21-OCT-1999;	99US-0160770P.
PR	21-OCT-1999;	99US-0160814P.
PR	21-OCT-1999;	99US-0160815P.
PR	22-OCT-1999;	99US-0160980P.

PR	22-OCT-1999;	99US-0160981P.	
PR	22-OCT-1999;	99US-0160989P.	
PR	25-OCT-1999;	99US-0161404P.	
PR	25-OCT-1999;	99US-0161405P.	
PR	25-OCT-1999;	99US-0161406P.	
PR	26-OCT-1999;	99US-0161359P.	
PR	26-OCT-1999;	99US-0161360P.	
PR	26-OCT-1999;	99US-0161361P.	
PR	28-OCT-1999;	99US-0161920P.	
PR	28-OCT-1999;	99US-0161922P.	
PR	28-OCT-1999;	99US-0161933P.	
PR	29-OCT-1999;	99US-0162142P.	
Query Match 27.6%; Score 298.8; DB 3; Length 2660;			
Best Local Similarity 59.0%; Pred. No. 2.2e-74;			
Matches 532; Conservative 0; Mismatches 367; Indels 3; Gaps 1;			
QY	92	TATACATGACACGAGGAGAAACAGTAATAGGCTAGCAACGGTAGTTCACACTTGTGGCT	151
DB	264	TTTACTATGATGAAGGTGGAAACAGTAATAGGAGCAACTCTGTTCAAACGGCTGCTT	323
QY	152	TTCTCTGTGCTACTTCCATATATATCTGTGTCATTTAAACACATGCACAACTGATAGAG	211
DB	324	TTCTATATCTATATCCCGCTT---TTCCTACTTCCGCTTCGGCTAGTAGTCTTT	380
QY	212	ATGGAAAAAGAACTCACCTAGGAACCGTGTATTTGGTTTACGTAGTGTGGACTTCTTTG	271
DB	381	CGGAGAGTTCGTGTTCACTCAAGTACATTTGTTGATCTATGTTTTCGTCGGTGTGATCA	440
QY	272	TAGGACAGATTGCTATCTGTACTCCATTTGGACTCTTTTACTTACCGGTTTCTACCTATT	331
DB	441	TTGCTGGAGATAATATGTTATATCTCTGTTGGACTTTTGTACCTCTCTGCATCGACGTATT	500
QY	332	CCCTGATCTGTGCATCTCAGTTAGCTTCAATGCTTTTCTCTCTTATTTTCTTAACCTAC	391
DB	501	CGCTCATTTGCGCTACTCAGTTAGCTTTCAACCGGTGTCTTTATTTTCATCAATGCTC	560
QY	392	AAAACTTACCCCTATCATTTTAAATTTCTTTTCTCTTAACTATATCTTCACCCCTAC	451
DB	561	AGAAGTTCACTGTTTGATTTCACTCCGTTGTTCTCTTTGTTCTCCGCTGCTTTGA	620
QY	452	TTGCATTCAATATAGGAGACAGACTCCACAAAAGTTACAAAAGAGAGATGTCAAG	511
DB	621	TAGCTCTCAATGATGATCGGATCTCTTCTGTGTCTCCAGGTCTAAGTATATTTGTTG	680
QY	512	GTTCATATGACCGTTGCTGCTGCTGCTGTTATGTTATGTTCTATGTTTATCCCTACACAGC	571
DB	681	GGTTTGTGTACACTTGTGCTGCTGCTCTCTATTTCTGTTGCTATCTTTATGCAAT	740
QY	572	TAGCCTTTCTAAAGTCTTAAAGAACAAATTTCTCAGAAGTTATGGATATGATAATCT	631
DB	741	TCTCGTTCGAGAAGATTCTGAGAGGGAGACGTTTCTGTGTTCTTGAATGCAATCT	800
QY	632	ACGTGAGTCTAGTGGCAGTTGTGTTAGCGTGTGGGCTTTTGTGCTAGCAGTGTGA	691
DB	801	ACACTTCTTTAGTGGCGACTCTGTGTTATAGGCTTTTGTAGCGGGAATGA	860
QY	692	AAACTTTGAGCAGTGAATGGAATACTACAAACATGGGAAGTATCTCATTTATCAACC	751
DB	861	GAACGCTGATGGGAATGGAAGGTTATCATAAAGGGCAAGCCTCTTATGTAAGTACCT	920
QY	752	TAGTGTGGACAGCTGTTACCTGGCAGTTATTTCTCCATCCGTTGGCAGGACTGATCTTCG	811
DB	921	TGCTCTGGACACAGTTACTTGGCAGTGTGTTCTGTTGGAGTCTGGGTTGATATTTTC	980
QY	812	AGCTCTCTCTCTATTTCTCAATGCAATAGCGTTTGGGATCCCGAGTGTTCCTATCT	871
DB	981	TGCTGACGTCTCTCTCTCAACGCTCATTTAGTACGCTCTCTCTAGCTGTGATCCACTAG	1040
QY	872	TGCTGCTAATCATTTTCCATGACAAATGAATGGGTTAAAGGTTATTTATGATCCCTAG	931
DB	1041	CAGCTTTGTTGTTTCCGTGATAAATGAGTGGTGTAAAGATTATATGGAATGCTGATCG	1100

QY	932	CTATTGGGGTTTCACTTCTCTATGCTTACCAACAATATCTTGTATGACAAAAAATTGAAGA	991
DB	1101	CTATTGGGGTTTTCGCTTCTTATGTTTACCAGAATCATATTTGATGACTTGAAGTAAGAC	1160
QY	992	AA 993	
DB	1161	AA 1162	
RESULT 10			
AAA97921 standard; DNA; 1145 BP.			
XX	AC	AAA97921;	
XX	DT	19-JAN-2001 (first entry)	
XX	DE	A. thaliana PUP1 DNA #3.	
XX	KW	PUP1; transgenic plant; nucleobase transporter; apical dominance;	
XX	KW	flowering behaviour; senescence; pesticide distribution; ds.	
XX	OS	Arabidopsis thaliana.	
XX	PN	DE19907209-A1.	
XX	PD	24-AUG-2000.	
XX	PF	19-FEB-1999; 99DE-01007209.	
XX	PR	19-FEB-1999; 99DE-01007209.	
XX	PA	(FROM/) FROMMER W.	
XX	PI	Gillissen B, Buerkle L, Andre B, Frommer WB;	
XX	XX	WPI; 2000-566202/53.	
XX	PT	Nucleic acid, useful for producing transgenic plants with altered	
XX	PT	nucleobase transport, encodes a nucleobase transporter protein of	
XX	XX	Arabidopsis thaliana.	
XX	PS	Claim 1f; Page 12-13; 24pp; German.	
XX	CC	This invention describes a novel nucleic acid encoding a plant nucleobase	
XX	CC	transporter (I). (I) is produced by complementation of a nucleobase	
XX	CC	transport (NBT)-defective host cell with a plant gene bank by selection	
XX	CC	of NBT-positive cells. (I) is used to isolate homologous sequences from	
XX	CC	bacteria, fungi, plants, animals and humans, for expression of the	
XX	CC	encoded protein (II) in prokaryotic or eukaryotic cells, for inhibiting	
XX	CC	expression of (II) (when in antisense orientation), and to produce	
XX	CC	transgenic crop plants. The transgenic plants have modified nucleobase	
XX	CC	transport properties, e.g. altered affinity and substrate specificity	
XX	CC	that may result in more efficient nucleobase transport in leaves, changes	
XX	CC	in apical dominance, flowering behaviour and senescence, or improved	
XX	CC	distribution of pesticides. This sequence encodes the Arabidopsis	
XX	CC	thaliana PUP1 protein which is described in the method of the invention	
XX	SQ	Sequence 1145 BP; 309 A; 234 C; 226 G; 376 T; 0 U; 0 Other;	
Query Match 27.6%; Score 298.4; DB 3; Length 1145;			
Best Local Similarity 58.3%; Pred. No. 2e-74;			
Matches 571; Conservative 0; Mismatches 391; Indels 18; Gaps 2;			
QY	89	GAGTATACCTATGACAAACGAGGAAACAGTAATAGCTAGCAACGGTAGTTCAACTGTG	148
DB	140	GACTTTACTAGCAAAAAGCGGTAAAGCACATGCTCGAAACCTTGGTTGAGTTGAG	199
QY	149	GCTTTCTCTGCTACTTCCATATATATATCTTGTCTATTTAAACACATCAACACTGATA	208
DB	200	GGTTTCTTTTAAACCTCTCTTGTCTATTTACTTTAAAGCCTGAGCGCTCAAGACTAAA	259
QY	209	GAGATGGAAAAAGAACCTCACCTAG-----GAACCGTGTATTGTTTACGTTAGTCTTG	262

Db 260 CAAATTACAAAATACTACTTCTCCCTCTTGACACTATCTTTAGTGATATATGGACTTG 319  
Qy 263 GACTTCTTGTAGGACAGATTGCTATCTGTACTCCTATTGGACTTCTTTTACTTACCCGTTT 322  
Db 320 GCTTGTCTTGTCTGCTGGACATGATATTTGTACTCATTTTGGGCTACTTTACCTTCTGTCT 379  
Qy 323 CTACCTATTCCCTGATCTGTCATCTCAGTTAGCCTTCAATGCTTTCTTCTCTTATTTTC 382  
Db 380 CAACCTTCTCTTTGATCTCTGGTCGCAATTTGGCTTTTAACGCGTCTTCTTACTTCC 439  
Qy 383 TTAACCTACAAAATACTACCCCTATCATTTTAAATTCCTTTTCTCTTAACTATATCTT 442  
Db 440 TAAACTCACAAAATAACACCAATTTATCACTTCACTTGTCTCTTACCATACTT 499  
Qy 443 CCACCTACTTGCATTCATTAATATGAGGAGACAGACTCCACAAAAGTTACAAAAGAG--- 499  
Db 500 CTACACTTCTTGTATCAACATGAACCAAGATCTCCCTCTTCTACTTCAAAAGTCCGAC 559  
Qy 500 -----AGTATGTCAAAGTTTCATATGCACCGTTGCTCGCTCTGCTGTTATGTC 550  
Db 560 CCAAGTCCAAGTATGTGATGGATACATCTCGCGGTCGGTAGCTCAGCTGGTTATTCTC 619  
Qy 551 TAGTCTTATCCCTACAAACAGCTAGCTTTCTAAAGTCTTAAAGAGCAAAATTTCTCAG 610  
Db 620 TGGTCTTCTTTAAACAGATTACGGTTTCGAAAGATTTCTAAGAAATACACATTCAGG 679  
Qy 611 AAGTTATGATATGATATCTACGTGAGTCTAGTGGCCAGTTGTGTAGCGTGTGGGGC 670  
Db 680 CTAATTTAGACATGGCCACATATCGTCTATGTGTAGCTACTTGTGTAGTTGGTAGGAC 739  
Qy 671 TTTTGTAGCAGTACGAGTGGAACTTTGACGAGTGAATGGAATACATACTACAAACATGGA 730  
Db 740 TTTTGGAAAGTGGTGGGGAAGCTGAGTACAGAAATGGAAGATTTCAACTAGGGA 799  
Qy 731 AGGTATCTACATATGATACCTAGTGTGGACAGCTGTTTACCTGGCAGTTATTTCTCCATCG 790  
Db 800 AAAGCTCATACATTTTGATAACATCGTTCAACGATATCATGGCAAGCTTGTGTGATTG 859  
Qy 791 GTGGCACAGGACTGATCTTCGAGCTCTCTCTCTATCTTCAATTCAAATGCAATAGCGTTTGG 850  
Db 860 GAAAGTGTGGTTGATATCGAAGTTTCATCGCTTTTTCCTCAATGTCTAAGCACTCTTT 919  
Qy 851 GACTCCCACTGGTCTCTATCTTGGCTGAATCATTTTCCATGACAAATGAATGGTTAA 910  
Db 920 GTTTACCAAGTGTGCGCTGTTCTGTCTGTCTCTCCGTGATGAGTGAATCA 979  
Qy 911 AGGTGATTTCTATGATCCTAGCTATTTGGGGTTTCACTTCTCTATGCTTACCACAAATATC 970  
Db 980 AGTTGGTTGCAATGTTTTTGGCCATCTGGGATTTGTTTCTTATGGTTATCAGCATATG 1039  
Qy 971 TTGATGACAAAACCTTGAAGAAAATATGAAATCAACAAACAGAAATCCCTGACCCAC 1030  
Db 1040 TCAATGATAGAAAGCCAGAAGAACCAAGAGCTTCTCTAGTCTTAAAGAAGAAGAAC 1099  
Qy 1031 CAGAGCAGAGAGTCAACT 1050  
Db 1100 AAAACAAGTAGATACCAATT 1119

RESULT 11  
ID AAC51119 standard; DNA; 1292 BP.  
XX  
AC AAC51119;  
XX  
DT 18-OCT-2000 (first entry)  
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 67344.  
XX Hybridisation assay; genetic mapping; gene expression control;  
KW protein identification; signal transduction pathway; metabolic pathway;  
KW promoter; termination sequence; ss.

XX OS Arabidopsis thaliana.  
XX  
FN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-00301439.  
XX  
PR 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 28-APR-1999; 99US-0130891P.  
PR 30-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 23-JUN-1999; 99US-0140355P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.



PR	29-JUN-1999;	99US-0140991P.
PR	30-JUN-1999;	99US-0141287P.
PR	01-JUL-1999;	99US-0141842P.
PR	01-JUL-1999;	99US-0142154P.
PR	02-JUL-1999;	99US-0142055P.
PR	06-JUL-1999;	99US-0142390P.
PR	08-JUL-1999;	99US-0142803P.
PR	09-JUL-1999;	99US-0142920P.
PR	12-JUL-1999;	99US-0142977P.
PR	13-JUL-1999;	99US-0143342P.
PR	14-JUL-1999;	99US-0143624P.
PR	15-JUL-1999;	99US-0144005P.
PR	16-JUL-1999;	99US-0144085P.
PR	16-JUL-1999;	99US-0144086P.
PR	18-JUL-1999;	99US-0144325P.
PR	19-JUL-1999;	99US-0144331P.
PR	19-JUL-1999;	99US-0144332P.
PR	19-JUL-1999;	99US-0144333P.
PR	19-JUL-1999;	99US-0144334P.
PR	19-JUL-1999;	99US-0144335P.
PR	20-JUL-1999;	99US-0144352P.
PR	20-JUL-1999;	99US-0144632P.
PR	20-JUL-1999;	99US-0144884P.
PR	21-JUL-1999;	99US-0144814P.
PR	21-JUL-1999;	99US-0145086P.
PR	21-JUL-1999;	99US-0145088P.
PR	22-JUL-1999;	99US-0145087P.
PR	22-JUL-1999;	99US-0145089P.
PR	22-JUL-1999;	99US-0145192P.
PR	23-JUL-1999;	99US-0145145P.
PR	23-JUL-1999;	99US-0145218P.
PR	23-JUL-1999;	99US-0145224P.
PR	26-JUL-1999;	99US-0145276P.
PR	27-JUL-1999;	99US-0145913P.
PR	27-JUL-1999;	99US-0145918P.
PR	27-JUL-1999;	99US-0145919P.
PR	28-JUL-1999;	99US-0145951P.
PR	02-AUG-1999;	99US-0146386P.
PR	02-AUG-1999;	99US-0146388P.
PR	02-AUG-1999;	99US-0146389P.
PR	03-AUG-1999;	99US-0147038P.
PR	04-AUG-1999;	99US-0147204P.
PR	04-AUG-1999;	99US-0147302P.
PR	05-AUG-1999;	99US-0147192P.
PR	05-AUG-1999;	99US-0147260P.
PR	06-AUG-1999;	99US-0147303P.
PR	06-AUG-1999;	99US-0147416P.
PR	09-AUG-1999;	99US-0147493P.
PR	09-AUG-1999;	99US-0147935P.
PR	10-AUG-1999;	99US-0148171P.
PR	11-AUG-1999;	99US-0148319P.
PR	12-AUG-1999;	99US-0148341P.
PR	13-AUG-1999;	99US-0148565P.
PR	13-AUG-1999;	99US-0148684P.
PR	16-AUG-1999;	99US-0149368P.
PR	17-AUG-1999;	99US-0149175P.
PR	18-AUG-1999;	99US-0149426P.
PR	20-AUG-1999;	99US-0149722P.
PR	20-AUG-1999;	99US-0149723P.
PR	20-AUG-1999;	99US-0149929P.
PR	23-AUG-1999;	99US-0149902P.
PR	23-AUG-1999;	99US-0149930P.
PR	25-AUG-1999;	99US-0150566P.
PR	26-AUG-1999;	99US-0150884P.
PR	27-AUG-1999;	99US-0151065P.
PR	27-AUG-1999;	99US-0151066P.
PR	27-AUG-1999;	99US-0151080P.
PR	30-AUG-1999;	99US-0151303P.
PR	31-AUG-1999;	99US-0151438P.
PR	01-SEP-1999;	99US-0151930P.
PR	07-SEP-1999;	99US-0152363P.
PR	10-SEP-1999;	99US-0153070P.
PR	13-SEP-1999;	99US-0153758P.
PR	15-SEP-1999;	99US-0154018P.
PR	16-SEP-1999;	99US-0154039P.
PR	20-SEP-1999;	99US-0154779P.
PR	22-SEP-1999;	99US-0155139P.
PR	23-SEP-1999;	99US-0155486P.
PR	24-SEP-1999;	99US-0155659P.
PR	28-SEP-1999;	99US-0156458P.
PR	29-SEP-1999;	99US-0156596P.
PR	04-OCT-1999;	99US-0157117P.
PR	05-OCT-1999;	99US-0157753P.
PR	06-OCT-1999;	99US-0157865P.
PR	07-OCT-1999;	99US-0158029P.
PR	08-OCT-1999;	99US-0158232P.
PR	12-OCT-1999;	99US-0158369P.
PR	13-OCT-1999;	99US-0159293P.
PR	13-OCT-1999;	99US-0159294P.
PR	13-OCT-1999;	99US-0159295P.
PR	14-OCT-1999;	99US-0159329P.
PR	14-OCT-1999;	99US-0159330P.
PR	14-OCT-1999;	99US-0159331P.
PR	14-OCT-1999;	99US-0159637P.
PR	14-OCT-1999;	99US-0159638P.
PR	18-OCT-1999;	99US-0159584P.
PR	21-OCT-1999;	99US-0160741P.
PR	21-OCT-1999;	99US-0160767P.
PR	21-OCT-1999;	99US-0160768P.
PR	21-OCT-1999;	99US-0160770P.
PR	21-OCT-1999;	99US-0160814P.
PR	21-OCT-1999;	99US-0160815P.
PR	22-OCT-1999;	99US-0160980P.
PR	22-OCT-1999;	99US-0160981P.
PR	22-OCT-1999;	99US-0160989P.
PR	25-OCT-1999;	99US-0161404P.
PR	25-OCT-1999;	99US-0161405P.
PR	25-OCT-1999;	99US-0161406P.
PR	26-OCT-1999;	99US-0161359P.
PR	26-OCT-1999;	99US-0161360P.
PR	26-OCT-1999;	99US-0161361P.
PR	28-OCT-1999;	99US-0161920P.
PR	28-OCT-1999;	99US-0161992P.
PR	28-OCT-1999;	99US-0161993P.
PR	29-OCT-1999;	99US-0162142P.
<hr/>		
Query Match 27.6%; Score 298.4; DB 3; Length 1292;		
Best Local Similarity 58.3%; Pred. No. 2,1e-74;		
Matches 571; Conservative 0; Mismatches 391; Indels 18; Gaps 2;		
QY	89	GAGTATACATGACAAACGGAGGAAACAGTAAATGGCTAGCAACGGTAGTTCACCTGTTG 148
Db	191	GACTTTACTACGAAAAAGCGGTAAAGCACATGGCTCGAAACCTTGGTTTCAGCTTGTAG 250
QY	149	GCTTTCTGTGCTACTTCCATATTATATCTGTCAITTTAAACACATGCAACACTGATA 208
Db	251	GGTTTCTTTTAAACCTTCTTGTCTATTATTACTTAAAGCCTGAGCGCTCAAGACTTAAA 310
QY	209	GAGATGGAAAAAGAACCTCACCTAG-----GAAACCGTGTATTGGTTTACGTAGTCTTG 262
Db	311	CCATTACCAAAAAAACTACTTCTTCTTCTTGACATATCTTTAGTATATTGGAATG 370
QY	263	GACTTCTTTGTAGGACAGATTGCTATCTGTACTCCATTGGACTTCTTTACTTACCGTTT 322
Db	371	GCTTGTCTTGTGCTGGACATTTTGTACTCATTTTGGGCTACTTTTACCTTCTCTGTCT 430
QY	323	CTACCTATTCCTGTGATCTGTGATCTCAGTTAGCTTCAATGCTTCTTCTTATTTTC 382
Db	431	CAACTTCTCTTTGTATCTCTGCTCGCAATTGGCTTTTAAACCGCTCTTCTCTTACTCC 490
QY	383	TTAACTCAGAAAACTTACCCCTATCATTTTAAATTTCTTTTCTTCTTAACTATATCTT 442
Db	491	TAACTCAGAAAAATCACACATTTACTCAATTCATCTGTTCTCTTAAACATATCTT 550
QY	443	CCACCTTACTTGTCAITTTCAATTAATGAGGAGACAGACTCCACAAAAAGTTACAAAAGGAG--- 499



QY	812	AGCTCTCCTCTCTATCTCAATGCAATAGCGTTTGGGACTCCAGTGGTTCCCTATCT	871	PR	28-MAY-1999;	99US-0136782P.
Db	68018	TGGTGCTTCGCTGTTTCAAAACCTTATTAGTAGCGTCTCCTCACTCATTTGACGCTCTCG	68077	PR	01-JUN-1999;	99US-0137222P.
QY	872	TGCGCTGTAATCATTTTCCATGACAAATGAATGGGTTAAAGGTGATTTCCTATGATCCTAG	931	PR	03-JUN-1999;	99US-0137528P.
Db	68078	CGGCCATTGGCGTGTTCATGACAGCTGACTGAGGTTAGATGGTGGCGATCCCATCG	68137	PR	04-JUN-1999;	99US-0137502P.
QY	932	CTATTTGGGGTTTCACTTCCTATGTCTACCAACAATATCTTGATGACAAAACTTGAAGA	991	PR	07-JUN-1999;	99US-0137722P.
Db	68138	CCTTCACAGGATTCACGCTTTTATATCTACCAAGTACTCTTGATGACTTGAAAGTACAA	68197	PR	08-JUN-1999;	99US-0138094P.
QY	992	AAATCATGAA 1002		PR	10-JUN-1999;	99US-0138540P.
Db	68198	GAGCACGAGAA 68208		PR	14-JUN-1999;	99US-0138847P.
				PR	16-JUN-1999;	99US-0139119P.
				PR	17-JUN-1999;	99US-0139453P.
				PR	18-JUN-1999;	99US-0139454P.
				PR	18-JUN-1999;	99US-0139455P.
				PR	18-JUN-1999;	99US-0139456P.
				PR	18-JUN-1999;	99US-0139457P.
				PR	18-JUN-1999;	99US-0139458P.
				PR	18-JUN-1999;	99US-0139459P.
				PR	18-JUN-1999;	99US-0139460P.
				PR	18-JUN-1999;	99US-0139461P.
				PR	18-JUN-1999;	99US-0139462P.
				PR	18-JUN-1999;	99US-0139463P.
				PR	18-JUN-1999;	99US-0139750P.
				PR	18-JUN-1999;	99US-0139763P.
				PR	21-JUN-1999;	99US-0139817P.
				PR	22-JUN-1999;	99US-0139899P.
				PR	23-JUN-1999;	99US-0140353P.
				PR	23-JUN-1999;	99US-0140354P.
				PR	24-JUN-1999;	99US-0140695P.
				PR	28-JUN-1999;	99US-0140823P.
				PR	29-JUN-1999;	99US-0140991P.
				PR	30-JUN-1999;	99US-0141287P.
				PR	01-JUL-1999;	99US-0141842P.
				PR	01-JUL-1999;	99US-0142154P.
				PR	02-JUL-1999;	99US-0142055P.
				PR	06-JUL-1999;	99US-0142390P.
				PR	08-JUL-1999;	99US-0142803P.
				PR	09-JUL-1999;	99US-0142920P.
				PR	12-JUL-1999;	99US-0142977P.
				PR	13-JUL-1999;	99US-0143542P.
				PR	14-JUL-1999;	99US-0143624P.
				PR	15-JUL-1999;	99US-0144003P.
				PR	16-JUL-1999;	99US-0144085P.
				PR	16-JUL-1999;	99US-0144086P.
				PR	19-JUL-1999;	99US-0144325P.
				PR	19-JUL-1999;	99US-0144331P.
				PR	19-JUL-1999;	99US-0144332P.
				PR	19-JUL-1999;	99US-0144333P.
				PR	19-JUL-1999;	99US-0144334P.
				PR	19-JUL-1999;	99US-0144335P.
				PR	20-JUL-1999;	99US-0144352P.
				PR	20-JUL-1999;	99US-0144632P.
				PR	20-JUL-1999;	99US-0144884P.
				PR	21-JUL-1999;	99US-0144814P.
				PR	21-JUL-1999;	99US-0145086P.
				PR	21-JUL-1999;	99US-0145088P.
				PR	22-JUL-1999;	99US-0145085P.
				PR	22-JUL-1999;	99US-0145087P.
				PR	22-JUL-1999;	99US-0145089P.
				PR	22-JUL-1999;	99US-0145192P.
				PR	23-JUL-1999;	99US-0145145P.
				PR	23-JUL-1999;	99US-0145218P.
				PR	23-JUL-1999;	99US-0145224P.
				PR	26-JUL-1999;	99US-0145276P.
				PR	27-JUL-1999;	99US-0145913P.
				PR	27-JUL-1999;	99US-0145918P.
				PR	28-JUL-1999;	99US-0145919P.
				PR	02-AUG-1999;	99US-0145951P.
				PR	02-AUG-1999;	99US-0146386P.
				PR	02-AUG-1999;	99US-0146388P.
				PR	02-AUG-1999;	99US-0146389P.
				PR	03-AUG-1999;	99US-0147038P.
				PR	04-AUG-1999;	99US-0147204P.

RESULT 13

AAC37652

ID AAC37652 standard; DNA; 1403 BP.

XX AAC37652;

XX 17-OCT-2000 (first entry)

XX Arabidopsis thaliana DNA fragment SEQ ID NO: 18172.

XX Hybridisation assay; genetic mapping; gene expression control;  
KW protein identification; signal transduction pathway; metabolic pathway;  
KW promoter; termination sequence; ss.

XX Arabidopsis thaliana.

XX EP1033405-A2.

XX 06-SEP-2000.

PF 25-FEB-2000; 2000EP-00301439.

XX 25-FEB-1999; 99US-0121825P.

PR 05-MAR-1999; 99US-0123180P.

PR 09-MAR-1999; 99US-0123548P.

PR 23-MAR-1999; 99US-0125788P.

PR 25-MAR-1999; 99US-0126264P.

PR 29-MAR-1999; 99US-0126785P.

PR 01-APR-1999; 99US-0127462P.

PR 06-APR-1999; 99US-0128234P.

PR 08-APR-1999; 99US-0128714P.

PR 16-APR-1999; 99US-0129845P.

PR 19-APR-1999; 99US-0130077P.

PR 21-APR-1999; 99US-0130449P.

PR 23-APR-1999; 99US-0130510P.

PR 23-APR-1999; 99US-0130891P.

PR 28-APR-1999; 99US-0131449P.

PR 30-APR-1999; 99US-0132048P.

PR 30-APR-1999; 99US-0132407P.

PR 04-MAY-1999; 99US-0132484P.

PR 05-MAY-1999; 99US-0132485P.

PR 06-MAY-1999; 99US-0132486P.

PR 06-MAY-1999; 99US-0132487P.

PR 07-MAY-1999; 99US-0132863P.

PR 11-MAY-1999; 99US-0134256P.

PR 14-MAY-1999; 99US-0134218P.

PR 14-MAY-1999; 99US-0134219P.

PR 14-MAY-1999; 99US-0134221P.

PR 14-MAY-1999; 99US-0134370P.

PR 18-MAY-1999; 99US-0134768P.

PR 19-MAY-1999; 99US-0134941P.

PR 20-MAY-1999; 99US-0135124P.

PR 21-MAY-1999; 99US-0135353P.

PR 24-MAY-1999; 99US-0135629P.

PR 25-MAY-1999; 99US-0136021P.

PR 27-MAY-1999; 99US-0136392P.

PR	04-AUG-1999;	99US-0147302P.
PR	05-AUG-1999;	99US-0147192P.
PR	05-AUG-1999;	99US-0147260P.
PR	06-AUG-1999;	99US-0147303P.
PR	06-AUG-1999;	99US-0147416P.
PR	09-AUG-1999;	99US-0147493P.
PR	09-AUG-1999;	99US-0147935P.
PR	10-AUG-1999;	99US-0148171P.
PR	11-AUG-1999;	99US-0148319P.
PR	12-AUG-1999;	99US-0148341P.
PR	13-AUG-1999;	99US-0148565P.
PR	13-AUG-1999;	99US-0148684P.
PR	16-AUG-1999;	99US-0149368P.
PR	17-AUG-1999;	99US-0149175P.
PR	18-AUG-1999;	99US-0149426P.
PR	20-AUG-1999;	99US-0149722P.
PR	20-AUG-1999;	99US-0149723P.
PR	20-AUG-1999;	99US-0149929P.
PR	23-AUG-1999;	99US-0149902P.
PR	23-AUG-1999;	99US-0149930P.
PR	25-AUG-1999;	99US-0150566P.
PR	26-AUG-1999;	99US-0150884P.
PR	27-AUG-1999;	99US-0151065P.
PR	27-AUG-1999;	99US-0151066P.
PR	27-AUG-1999;	99US-0151080P.
PR	30-AUG-1999;	99US-0151303P.
PR	31-AUG-1999;	99US-015138P.
PR	01-SEP-1999;	99US-0151930P.
PR	07-SEP-1999;	99US-0152363P.
PR	10-SEP-1999;	99US-0153070P.
PR	13-SEP-1999;	99US-0153758P.
PR	15-SEP-1999;	99US-0154018P.
PR	16-SEP-1999;	99US-0154039P.
PR	20-SEP-1999;	99US-0154779P.
PR	22-SEP-1999;	99US-0155139P.
PR	23-SEP-1999;	99US-0155486P.
PR	24-SEP-1999;	99US-0155659P.
PR	28-SEP-1999;	99US-0156458P.
PR	29-SEP-1999;	99US-0156596P.
PR	04-OCT-1999;	99US-0157117P.
PR	05-OCT-1999;	99US-0157753P.
PR	06-OCT-1999;	99US-0157865P.
PR	07-OCT-1999;	99US-0158029P.
PR	08-OCT-1999;	99US-0158232P.
PR	12-OCT-1999;	99US-0158369P.
PR	13-OCT-1999;	99US-0159293P.
PR	13-OCT-1999;	99US-0159294P.
PR	13-OCT-1999;	99US-0159295P.
PR	14-OCT-1999;	99US-0159329P.
PR	14-OCT-1999;	99US-0159330P.
PR	14-OCT-1999;	99US-0159331P.
PR	14-OCT-1999;	99US-0159637P.
PR	18-OCT-1999;	99US-0159638P.
PR	18-OCT-1999;	99US-0159584P.
PR	21-OCT-1999;	99US-0160741P.
PR	21-OCT-1999;	99US-0160767P.
PR	21-OCT-1999;	99US-0160768P.
PR	21-OCT-1999;	99US-0160770P.
PR	21-OCT-1999;	99US-0160814P.
PR	21-OCT-1999;	99US-0160815P.
PR	22-OCT-1999;	99US-0160980P.
PR	22-OCT-1999;	99US-0160981P.
PR	22-OCT-1999;	99US-0160989P.
PR	25-OCT-1999;	99US-0161404P.
PR	25-OCT-1999;	99US-0161405P.
PR	25-OCT-1999;	99US-0161406P.
PR	26-OCT-1999;	99US-0161359P.
PR	26-OCT-1999;	99US-0161360P.
PR	26-OCT-1999;	99US-0161361P.
PR	28-OCT-1999;	99US-0161920P.
PR	28-OCT-1999;	99US-0161992P.
PR	28-OCT-1999;	99US-0161993P.
PR	29-OCT-1999;	99US-0162142P.

Query Match

20.9%; Score 225.8; DB 3; Length 1403;

Best Local Similarity 54.0%; Pred. No. 1.1e-53;

Mismatches 492; Conservative 0; Mismatches 407; Indels 12; Gaps 1;

Qy	92	TATACTATGACAAAGGAGGAAACAGTAAATGGCTAGCAACGGTAGTTCAACTTCTGGCT	151
Db	345	TTTATTACAAGAGGTGGAAACAGTAAATGGATCTCTACTCTTGTCCAACTTGGT	404
Qy	152	TTCCGTGTCTACTTCCATATTATATCTTGTTCATTTAAACACATGCAACCACTGATAGAG	211
Db	405	TTCCGATTCTCTATCTCCCTCTTGTTCCTCTCTGCTCCACACTCTCTTCATCTCTT	464
Qy	212	ATGGAAGAAACCTCACCTAGGAACCGTGTATTTGGTTTACGTAGTCTGGACTTTTG	271
Db	465	GT-----TCTTCAAGACTCTGGTTTGGATTATCTTCTCTGGTTTCTGA	512
Qy	272	TAGGAGCAGATTGCTATCTGTACTCTCCATTGAGCTTCTTTACTTACCCGTTTCTACCTATT	331
Db	513	TTGGTTTGAACAATCTTTTATACCTTTTGGTCTTTTGTATCTCTCTGCTTCACTTATT	572
Qy	332	CCCTGATCTGTGCATCTCAGTTAGCCTTCAATGCTTCTTCTCTTATTTTCTTAACCTAC	391
Db	573	CGATTCTTTGTTCTTTCACAGTTAGCTTTCATGCTGCTCTCTTATTACATCAATTCTC	632
Qy	392	AAAACTTACCCCTATCAATTTTAAATCTCTTTTCTCTTAACTATATCTTCCACCCTAC	451
Db	633	AGAAATCACTGTTGTTGATTCTCTCTCTGTGTTGTTCTCTCTGTCTGTGTGTGG	692
Qy	452	TTGCATTCAATAATGAGGAGACAGATCCCAAAAGTTACAAAGGAGAGTATCTCAAAG	511
Db	693	TTTCTCTTGAGATGATTCCAATAGCCCATCAGAGATTCTAAGTGGAGTTACTTGTATTG	752
Qy	512	GTTCATATGACCGTTGCTGCGTCTGCTGGTTAGTGTCTAGTCTTATCCCTACAAACAGC	571
Db	753	GTTGTTTATGTACTGTTTTTGTCTTCTTATCTTCTCTTCTTCTTCTTCTCTTATGCACT	812
Qy	572	TAGCCTTTCTAAAAGTCTTAAAGAACAAATTTCTCAGAAAGTTATCGATGATGATATCT	631
Db	813	TTTCTTTTCGAGAAATGTTCTCAAGAGTGAGAGCTTCTCTATGGTTCTCGAGATGCAAACT	872
Qy	632	ACGTGAGTCTAGTGGCCAGTTGTTAGCGTGGTGGGCTTTTGTCTAGCAGTGAAGTGA	691
Db	873	ATAGTCTGCTTGTGGCTTCTTGTGTTGGGTTATCGGTTGTTGCGAGCGGGAATGA	932
Qy	692	AAACTTTGACGAGTGAATAAGATAACTTACAAACATGGAAGGATCTCTACATTAAGAAC	751
Db	933	TGTTGTTGAGTGTGGAGATGGAAGAGTTTCACGAGGTCAGTCAATTTATGTTTGACTT	992
Qy	752	TAGTGTGACAGCTGTTTACCTGGCAGTTATCTCCATCGGTGGCACAGGACTGATCTCG	811
Db	993	TGGTTGGGACCGCGGTTTCTGTCGCAATTTGGGTTCTGTAGGAGCGGTGCGGCTTATATTC	1052
Qy	812	AGCTCTCTCTCTATTCTCAAAATGCAATAAGCGTTTTTGGGACTCCCACTGGTCTCTATCT	871
Db	1053	TGGTTTCTTCTGCTGTTTTCGNACTTATTTGGTACACTCTCACCTATTTGACGCTCTTG	1112
Qy	872	TGGCTGTAATCAATTTCCATGACAAAATGAATGAGTGTAAAGGTGATTTCTATGATCTTAG	931
Db	1113	CAGCCATTGGCGTGTTCACGACAAAGCTGACTGAGGTTAAGATGGTCCGATGCTCATTTG	1172
Qy	932	CTATTGGGGTTTCACTTCTCTATGCTTACCACAAATATCTTGATGACAAAACCTTGAGA	991
Db	1173	CTTTCATGGGGTTTCGGGTTTTTATATCTACCAGAAATTTATCTTGATGACCTGAAAGTACAAA	1232
Qy	992	AAATCATGAA 1002	
Db	1233	GAGCAGAGAA 1243	

RESULT 14

ADA70748

ID ADA70748 standard; DNA; 2175 BP.

XX AC ADA70748;  
XX DT 20-NOV-2003 (first entry)  
XX DE Rice gene, SEQ ID 4071.  
XX KW Plant; bacterial infection; fungal infection; viral infection; rice;  
XX KW gene; ds.  
XX OS Oryza sativa.  
XX PN WO2003000898-A1.  
XX PD 03-JAN-2003.  
XX PF 22-JUN-2001; 2001WO-IB001105.  
XX PR 22-JUN-2001; 2001WO-IB001105.  
XX PA (SYGN ) SYNGENTA PARTICIPATIONS AG.  
XX PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;  
XX PI Katagiri F, Qian S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;  
XX DR WPI; 2003-175290/17.  
XX  
XX Identifying at least one gene involved in plant resistance or response to  
XX PT pathogenic infection for conferring resistance or tolerance to a plant to  
XX PT bacterial, fungal or viral infection by determining or detecting plant  
XX PT gene expression.  
XX  
XX Claim 6; SEQ ID NO 4071; 899pp; English.  
XX  
XX The present invention relates to a method (M1) for identifying genes  
XX CC involved in plant resistance or response to pathogenic infection. M1  
XX CC comprises identifying a gene whose expression is significantly altered in  
XX CC the incompatible interaction of plant gene expression relative to  
XX CC expression of the gene in an uninfected plant, in a mutant plant that  
XX CC does not express a gene associated with response to pathogenic infection,  
XX CC or in a corresponding incompatible or compatible interaction. (M1) is  
XX CC useful for conferring resistance to resistance or tolerance to a plant to  
XX CC bacterial, fungal or viral infection. The present sequence was used to  
XX CC illustrate the invention.  
XX  
XX Sequence 2175 BP; 363 A; 701 C; 671 G; 439 T; 0 U; 1 Other;  
XX  
XX Query Match 20.6%; Score 222.8; DB 8; Length 2175;  
XX Best Local Similarity 58.2%; Pred. No. 9.8e-53;  
XX Matches 392; Conservative 0; Mismatches 282; Indels 0; Gaps 0;  
XX  
QY 235 AACCGTGATTGTTACGTAGTGGACTTCTTGTAGGACGACATTGCTATCTGTAC 294  
DB 343 AAGCTGCCGTGATATACATCGTCTGGGCTCATCATGCCGCCGACGATGATGATC 402  
QY 295 TCATTGGACTCTTTACTTACCGCTTTCTACCTATTCCCTGATCTGTGCATCTCAGTTA 354  
DB 403 ACTGGTGGCTCAAGTACCTCCCGCTTCGACCTACTCGCTCATCTGTGCCACCGCTC 462  
QY 355 GCCTTCAAGCTTTCTTCTTATTTTCTTAACTCAAAAACCTTACCCCTATCATTTTA 414  
DB 463 GCGTTCAATGCTCTTCTCATACGTGCTCAACTCCGAGAGGCTCACTCTCTGATATTC 522  
QY 415 AATTCTCTTTCTTCTTAACTATATCTTCCACCTACTTGCATTAATGAGGAGACA 474  
DB 523 AACTCGCTGCTGCTGCTACCATGTCGGCTTCGCTCATCGAGTCAGAGGAGTCTCAG 582  
QY 475 GACTCCACAAAGTTACAAAGGAGAGATGTCAAAAGTTTTCATATGACCCGTTGCTGGG 534  
DB 583 GGGGTACCGGCTCTCGGAGGAGAGTATCTGCTCGGTTTCTGCTGACGCTGGGGCG 642  
QY 535 TCTGCTGGTATGGTCTAGTCTTATCCCTCAACACGCTAGCCCTTTCTAAAGTCTTAAAG 594

DB 643 TCGTGCACTACTCGCTGTGATCTCGCGCTGATGCACTCACCCTTCGAGACCATCATCAAG 702  
QY 595 AAGCAAAATTTCTCAGAAGTTATGATATGATAATCTAGCTGAGTCTAGTGGCCAGTTGT 654  
DB 703 AAGCACACCTTCTCAGCCGCTCCTCAACATGCAGATCTACACGGCGCTCGTGGCGACGGCC 762  
QY 655 GTTAGCGTGGTGGGCTTTTCTAGCAGTGAGTGAAAACCTTTGAGCAGTGAATGGAT 714  
DB 763 GCGTCCGTTGCTGGGCTGTTCCGAGCGCGAGTGGAGGTCGCTGAGGGGGAGATGAAC 822  
QY 715 AACTACAAACATGCGGAAGGTATCTTACATATGAACTAGTGTGGACAGCTGTACTCTG 774  
DB 823 GCGTTTCAGGTGCGGGCGAGTTCTCTACCTGATGACGCTGCTGTGGCGCGCGGTGTCGTGG 882  
QY 775 CAGTTATTCTCCATCGGTGGCACAGGACTGATCTTCGAGCTCTCTCTCTATTCTCAAAAT 834  
DB 883 CAGGTGGCCAAACATCGGGGTGCTCGGCTCATCTTCGAGGTGTCGGCGCTCTCTCTCAAC 942  
QY 835 GCAATAAGCGTTTGGGACTCCGAGTGGTTCCTATCTTGGCTGTAAATCAATTTTCCATGAC 894  
DB 943 GTGATCAGCACGCTGTCACTGCGCGTCAATCCCGTCTTCGCGTGGTGTGTTCCACGAC 1002  
QY 895 AAATGAATGGGTT 908  
DB 1003 AGGATGAACGGGT 1016  
DB  
RESULT 15  
ADA70213  
ID ADA70213 standard; DNA; 1152 BP.  
XX AC ADA70213;  
XX DT 20-NOV-2003 (first entry)  
XX DE Rice gene, SEQ ID 3536.  
XX KW Plant; bacterial infection; fungal infection; viral infection; rice;  
XX KW gene; ds.  
XX OS Oryza sativa.  
XX XX WO2003000898-A1.  
XX PN 03-JAN-2003.  
XX PD 22-JUN-2001; 2001WO-IB001105.  
XX PF 22-JUN-2001; 2001WO-IB001105.  
XX PR (SYGN ) SYNGENTA PARTICIPATIONS AG.  
XX PA Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;  
XX PI Katagiri F, Qian S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;  
XX DR WPI; 2003-175290/17.  
XX  
XX Identifying at least one gene involved in plant resistance or response to  
XX PT pathogenic infection for conferring resistance or tolerance to a plant to  
XX PT bacterial, fungal or viral infection by determining or detecting plant  
XX PT gene expression.  
XX  
XX Claim 6; SEQ ID NO 3536; 899pp; English.  
XX  
XX The present invention relates to a method (M1) for identifying genes  
XX CC involved in plant resistance or response to pathogenic infection. M1  
XX CC comprises identifying a gene whose expression is significantly altered in  
XX CC the incompatible interaction of plant gene expression relative to  
XX CC expression of the gene in an uninfected plant, in a mutant plant that  
XX CC does not express a gene associated with response to pathogenic infection,  
XX CC or in a corresponding incompatible or compatible interaction. (M1) is  
XX CC useful for conferring resistance to resistance or tolerance to a plant to  
XX CC bacterial, fungal or viral infection. The present sequence was used to  
XX CC illustrate the invention.

CC illustrate the invention.

```
XX
SQ Sequence 1152 BP; 294 A; 260 C; 248 G; 350 T; 0 U; 0 Other;

Query Match      19.6%; Score 211.6; DB 8; Length 1152;
Best Local Similarity 57.7%; Pred. No. 1.2e-49;
Matches 404; Conservative 0; Mismatches 284; Indels 12; Gaps 1;

QY 92 TATACATAGACACGAGGAAACAGTAAATGGCTAGCAACGGTAGTTCAACTTGTGGCT 151
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 329 TCTACTACAATCAAGCGGCAATAGCAAGTGGATGCCACATTCGCCAAACTGCTGGCT 388

QY 152 TTCC-----TGTGCTACTTCCATATATATATCTTGTCAITTTAAAAACACATGCAA 199
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 389 TTCCGATTTTGTTCATTCGCCCTATTTCTTTCCATTCAAAGACATCTTCTACACAACAG 448

QY 200 CAACGTAGATAGAGAAAGAAACCTCACCTAGGAACCGTGTATTTGGTTTACGTAGTGC 259
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 449 TCACCTAGTAGTCTCGCCCTACAATTTCTATCCCCAAAATTTACTCTGATATATGTTGCC 508

QY 260 TTGGACTTCTTGTAGGACGATTCGTATCTGTACTCCATTGGACTTCTTTACTTACCCG 319
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 509 TGGGCTCATCATTCCTCGACGACCTTGATGTATTCCTATGGCTACTATATCTTCCGG 568

QY 320 TTCTACCTATTTCCCTGATCTGTGCATCTCAGTTAGCCTTCAATGCTTTCTTCTCTTATT 379
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 569 TCTCAACATATTCGCTCATCTGTGCTAGTCACTTGCCTTCAATGCTGTCTTCTCATATT 628

QY 380 TTCTTAACTCACAAAACCTTACCCCTATCATTTTAAATTTCTTTTCTCTTAACTATAT 439
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 629 TCCTCAATGCTCAAAAATTCACCCCTCTGATTTTCAATTCGGTAGTCTCCTTACGTTTT 688

QY 440 CTTCCACCCTACTTGCATTCATATAGGAGACAGACTCCACAAAAGTTACAAAAGGAG 499
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 689 CTGCTTCACTCCTTGGAGTTGATGAAGATTCTCAGGGAACTACTAGTATATACATGGGA 748

QY 500 AGTATGTCAAAGGTTTCATATGCAACCGTTGCTGCTGTGCTGTATGGTCTAGTCTTTAT 559
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 749 AGTACATTTTGGGTTTCTGTTGACACTAGGGGCATCAGCTACATCTCGCTCATTTCTCT 808

QY 560 CCCTACAAACAGTAGCCCTTTCTAAAGTCCCTAAAGAGCAAAATTTCTCAGAAAGTTATGG 619
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 809 CCCTGATGCAAGTCACATTTGAGAAGGTTATTAAGAGGGAGACCTTCTCAGTTGTGTGA 868

QY 620 ATATGATAATCTACGTAGTCTAGTGGCCAGTTGTGTTAGCGTGTGGGGCTTTTTCGCTA 679
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 869 ACATGCAGATATATACAGCTCTCGTGGCAACATTTGGCTTCTCTTGTGGGTTATTTGCAA 928

QY 680 GCAGTGAGTGGAAAACCTTTGAGCAGTGAATGGATACTACAAAACATGGGAAGGTATCCT 739
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 929 GTGGTGAATGGATGACTTTTACAAGGAGAGATGCATTCACATTCAGTCTCGGAAGCTGTCA 988

QY 740 ACATTATGAACCTAGTGTGGACAGCTGTTACTTGGCAGTT 779
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 989 ATGTAATGACACTGCTGTGGCGGCTATATCTTTGGCAGAT 1028
```

Search completed: November 1, 2004, 07:20:05  
Job time : 571.648 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 1, 2004, 05:48:03 ; Search time 119.284 Seconds  
(without alignments)  
7704.726 Million cell updates/sec

Title: US-09-913-767-4

Perfect score: 1293

Sequence: 1 ggaagctctcttagagtgtg.....tatgatcaaacatatattcc 1293

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.\*

- 1: /cgn2\_6/prodata/1/ina/5A COMB.seq.\*
- 2: /cgn2\_6/prodata/1/ina/5B COMB.seq.\*
- 3: /cgn2\_6/prodata/1/ina/6A COMB.seq.\*
- 4: /cgn2\_6/prodata/1/ina/6B COMB.seq.\*
- 5: /cgn2\_6/prodata/1/ina/PCTUS COMB.seq.\*
- 6: /cgn2\_6/prodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	79.2	6.1	7218	1	US-08-232-463-14
2	44.2	3.4	1141	4	US-09-806-708B-22
3	42.8	3.3	399	4	US-09-621-976-8976
4	41.4	3.2	1141	4	US-09-806-708B-22
5	38.2	3.0	696	3	US-09-461-697-193
6	38.2	3.0	699	3	US-09-461-697-191
7	38.2	3.0	717	3	US-09-461-697-189
8	38.2	3.0	774	3	US-09-461-697-187
9	38.2	3.0	819	3	US-09-461-697-185
10	38.2	3.0	1669	3	US-09-461-697-184
11	38.2	3.0	1664976	4	US-08-916-421B-1
12	38.2	3.0	1664976	4	US-09-692-570-1
13	38.2	2.9	11917	4	US-09-566-921-32
14	35.6	2.8	264	4	US-09-107-532A-2112
15	35.6	2.8	600	4	US-09-107-532A-3027
16	35.6	2.8	832	4	US-09-621-976-2813
17	35.6	2.8	1047	3	US-09-198-955A-7
18	35.6	2.8	1047	3	US-09-694-531-7
19	35.6	2.8	1047	4	US-10-072-152-7
20	35.6	2.8	1266	4	US-09-248-796A-6306
21	35.2	2.7	504	3	US-09-134-001C-1508
22	35.2	2.7	1827	4	US-09-270-767-1308
23	35.2	2.7	1827	4	US-09-270-767-16590
24	35	2.7	312	4	US-09-621-976-18100
25	34.6	2.7	1394	3	US-09-247-155-76
26	34.4	2.7	474	4	US-09-621-976-18033
27	34.2	2.6	289	3	US-09-007-005-17

c	28	34.2	2.6	289	3	US-09-244-796-17	Sequence 17, Appl
	29	34.2	2.6	832	4	US-09-621-976-2813	Sequence 2813, Ap
	30	34.2	2.6	246240	2	US-08-724-394A-20	Sequence 20, Appl
	31	34.2	2.6	246240	2	US-08-724-394A-21	Sequence 21, Appl
	32	34.2	2.6	246240	2	US-08-724-394A-21	Sequence 22, Appl
	33	34	2.6	1794	4	US-09-248-796A-2292	Sequence 2292, Ap
	34	34	2.6	1923	4	US-09-134-000C-466	Sequence 1, Appl
	35	33.8	2.6	1332	4	US-09-642-894A-1	Sequence 107, App
	36	33.6	2.6	2072	4	US-09-774-528-107	Sequence 106, App
	37	33.6	2.6	2440	4	US-09-774-528-106	Sequence 4, Appli
	38	33.6	2.6	80246	3	US-09-078-294-4	Sequence 3, Appli
	39	33.6	2.6	80595	3	US-09-078-294-3	Sequence 3, Appli
	40	33.4	2.6	16550	4	US-08-916-421B-3	Sequence 3, Appli
	41	33.4	2.6	16550	4	US-09-692-570-3	Sequence 207, App
	42	33.2	2.6	231	3	US-09-461-697-207	Sequence 205, App
	43	33.2	2.6	282	3	US-09-461-697-205	Sequence 203, App
	44	33.2	2.6	306	3	US-09-461-697-203	Sequence 42, Appl
	45	33	2.6	1473	4	US-09-244-805-42	

ALIGNMENTS

RESULT 1  
US-08-232-463-14  
; Sequence 14, Application US/08232463  
; Patent No. 5670367  
; GENERAL INFORMATION:  
; APPLICANT: DORNER, F.  
; APPLICANT: SCHEIFLINGER, F.  
; APPLICANT: FALKNER, F. G.  
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Hardner  
; STREET: 1800 Diagonal Road, Suite 500  
; CITY: Alexandria  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22313-0299  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/232,463  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/935,313  
; FILING DATE:  
; APPLICATION NUMBER: EP 91 114 300.6  
; FILING DATE: 26-AUG-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BENT, Stephen A.  
; REGISTRATION NUMBER: 29,768  
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703)836-9300  
; TELEFAX: (703)683-4109  
; TELEX: 899149  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7218 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; CLONE: pTZ5pt-F1s  
US-08-232-463-14

Query Match 6.1%; Score 79.2; DB 1; Length 7218;







```
Qy 670 AAC 672
Db 315 ATC 313

RESULT 6
US-09-461-697-191/c
; Sequence 191, Application US/09461697
; Patent No. 6277974
; GENERAL INFORMATION:
; APPLICANT: COGENT NEUROSCIENCE, Inc.
; APPLICANT: Lo, Donald C.
; APPLICANT: Barney, Shawn
; APPLICANT: Thomas, Mary Beth
; APPLICANT: Portbury, Stuart D.
; APPLICANT: Puranam, Kasturi
; APPLICANT: Katz, Lawrence C.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING
; TITLE OF INVENTION: AND TREATING CONDITIONS, DISORDERS, OR DISEASES INVOLVING
; FILE REFERENCE: 10001-005-999
; CURRENT APPLICATION NUMBER: US/09/461.697
; CURRENT FILING DATE: 1999-12-14
; NUMBER OF SEQ ID NOS: 466
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 191
; LENGTH: 699
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-461-697-191

Query Match 3.0%; Score 38.2; DB 3; Length 699;
Best Local Similarity 47.3%; Pred. No. 0.08;
Matches 115; Conservative 0; Mismatches 128; Indels 0; Gaps 0;

Qy 430 TTCTCTTCCCTCACCAATCTTGGATCAGTTTACATCGTTACTGGACATATTAGTCTGCT 489
Db 558 TTCTCTCTCTTTTCACTCTTTGAGCTTTTACATCTTTCCCTCTTTTATCTCCCTT 499

Qy 490 AACTCTTATATGCTCTGTGGTTTACTACTATACCAAGTTTCTACTTTCTCCCTCATC 549
Db 498 CTCCTTTCCACTTCACTCTCTTTTCCAACTCCGTCTTTCTTCTGCTCTCTCTTT 439

Qy 550 TTGGCCTCACAAATGGCCCTTCACTGCTTCTTCTCAATATTTCTAACTCGCAGAGTTC 609
Db 438 TTCACTCTCTCTCTCTCTTTTCCATCTTCTCCATTTCTCTTCCGTTTCACTTTT 379

Qy 610 ACACCTTTCAATGTGAATCTCTGTCTTCTCTTACTATTTCTCTGCTCTGCTCTGCTGTC 669
Db 378 TCCATCTCTTCTCTTTTGGCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTTCC 319

Qy 670 AAC 672
Db 318 ATC 316

RESULT 7
US-09-461-697-189/c
; Sequence 189, Application US/09461697
; Patent No. 6277974
; GENERAL INFORMATION:
; APPLICANT: COGENT NEUROSCIENCE, Inc.
; APPLICANT: Lo, Donald C.
; APPLICANT: Barney, Shawn
; APPLICANT: Thomas, Mary Beth
; APPLICANT: Portbury, Stuart D.
; APPLICANT: Puranam, Kasturi
; APPLICANT: Katz, Lawrence C.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING
; TITLE OF INVENTION: AND TREATING CONDITIONS, DISORDERS, OR DISEASES INVOLVING
; FILE REFERENCE: 10001-005-999
```

```
; CURRENT APPLICATION NUMBER: US/09/461.697
; CURRENT FILING DATE: 1999-12-14
; NUMBER OF SEQ ID NOS: 466
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 189
; LENGTH: 717
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-461-697-189

Query Match 3.0%; Score 38.2; DB 3; Length 717;
Best Local Similarity 47.3%; Pred. No. 0.081;
Matches 115; Conservative 0; Mismatches 128; Indels 0; Gaps 0;

Qy 430 TTCTCTTCCCTCACCAATCTTGGATCAGTTTACATCGTTACTGGACATATTAGTCTGCT 489
Db 576 TTCTCTCTCTTTTCACTCTTTTGAATTTTACATCTTTCCCTCTTTTATCTCCCTT 517

Qy 490 AACTCTTATATGCTCTGTGGTTTACTACTATACCAAGTTTCTACTTTCTCCCTCATC 549
Db 516 CTCCTTTCCACTCTCACTCTCTTTTCCAACTCCGTCTTTCTTCTGCTCTCTCTTT 457

Qy 550 TTGGCCTCACAAATGGCCCTTCACTGCTCTCTCTCAATATTTCTAACTCGCAGAGTTC 609
Db 456 TTCACTCTCTCTCTCTCTTTTCCATCTTCTCCATTTCTCCGTCTTCACTTTT 397

Qy 610 ACACCTTTCAATGTGAATCTCTGTCTCTCTTACTATTTCTCTGCTCTGCTGCTGTC 669
Db 396 TCCATCTCTCTCTCTCTTTTGCATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTTCC 337

Qy 670 AAC 672
Db 336 ATC 334

RESULT 8
US-09-461-697-187/c
; Sequence 187, Application US/09461697
; Patent No. 6277974
; GENERAL INFORMATION:
; APPLICANT: COGENT NEUROSCIENCE, Inc.
; APPLICANT: Barney, Shawn
; APPLICANT: Thomas, Mary Beth
; APPLICANT: Portbury, Stuart D.
; APPLICANT: Puranam, Kasturi
; APPLICANT: Katz, Lawrence C.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING
; TITLE OF INVENTION: AND TREATING CONDITIONS, DISORDERS, OR DISEASES INVOLVING
; FILE REFERENCE: 10001-005-999
; CURRENT APPLICATION NUMBER: US/09/461.697
; CURRENT FILING DATE: 1999-12-14
; NUMBER OF SEQ ID NOS: 466
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 187
; LENGTH: 774
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-461-697-187

Query Match 3.0%; Score 38.2; DB 3; Length 774;
Best Local Similarity 47.3%; Pred. No. 0.085;
Matches 115; Conservative 0; Mismatches 128; Indels 0; Gaps 0;

Qy 430 TTCTCTTCCCTCACCAATCTTGGATCAGTTTACATCGTTACTGGACATATTAGTCTGCT 489
Db 633 TTCTCTCTCTTTTCACTCTTTTGAATTTTACATCTTTCCCTCTTTTATCTCCCTT 574

Qy 490 AACTCTTATATGCTCTGTGGTTTACTACTATACCAAGTTTCTACTTTCTCCCTCATC 549
Db 573 CTCCTTTCCACTCTCACTCTCTTTTCCAACTCTCTCTCTCTCTCTCTCTCTCTCTT 514
```







```
; NAME/KEY: misc feature
; LOCATION: (622708)..(622708)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (657081)..(657081)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (657203)..(657203)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (674435)..(674435)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (682442)..(682442)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (713652)..(713652)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (741684)..(741684)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (779455)..(779455)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (779676)..(779676)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (855539)..(855539)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (871619)..(871619)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1084830)..(1084830)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1096846)..(1096846)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1119881)..(1119881)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1130881)..(1130881)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1310988)..(1310988)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1313224)..(1313224)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1349473)..(1349473)
; OTHER INFORMATION: n equals a, t, c, or g
;
; Query Match 3.0%; Score 38.2; DB 4; Length 1664976;
```

```
Best Local Similarity 61.6%; Pred. No. 11;
Matches 61; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

Qy 1046 CCATAACTGCTGGGATGCGCTATAGTCCAGTTGAGCAGTATGTTTTCATGATA 1105
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 319372 CCATAAATTATGGTGAATATCTTTTCCCTCCAGATGCAGATGTTATTTTATAAAGTTG 319431

Qy 1106 AAATGACGCGTCAAGATCTTCTCCCATCATTTTAGCTA 1144
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 319432 AAAATATGATATAAAAAATAATAGCAATCAATTCAGTTA 319470

RESULT 13
US-09-566-921-32/c
; Sequence 32, Application US/09566921
; Patent No. 6682888
; GENERAL INFORMATION:
; APPLICANT: Loring, Jeanne P.
; APPLICANT: Tingley, Debora W.
; APPLICANT: Edwards, Carla M.
; TITLE OF INVENTION: GENES EXPRESSED IN ALZHEIMER'S DISEASE
; FILE REFERENCE: PA-0024 US
; CURRENT APPLICATION NUMBER: US/09/566,921
; CURRENT FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 138
; SOFTWARE: PERL Program
; SEQ ID NO 32
; LENGTH: 11917
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6682888 475473.1
; NAME/KEY: unsure
; LOCATION: 9384, 10290-10321, 10325
; OTHER INFORMATION: a, t, c, g, or other
US-09-566-921-32
```

```
Query Match 2.9%; Score 38; DB 4; Length 11917;
Best Local Similarity 52.2%; Pred. No. 0.56;
Matches 83; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

Qy 251 TCTTTGCTTGCTTGCAAGCACTTTCTACAATTTTGGCGAGAGTTTACTATGAAAATG 310
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 9509 TGTTAATCATGATCTGCTTCCAGGTTCAAGTTCATGATGACATGATGGAATTTT 9450

Qy 311 GTGGGAAGACTACATGATGGGAACACTTGTCCAACTAAATCGGCTTCCCTGTCTGTTTC 370
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 9449 CTCTAAACTCTACATGAACGCTAGGAGTTGAGAATTCATTTGTTTCAATGTAATAATTC 9390

Qy 371 TCTTCGCTTCTTTTCCAAACCAAAAATCCCAACCA 409
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 9389 TTTTANTTTTTTTTTTCCAGAAAGAAAATCCCAAGCAA 9351
```

```
RESULT 14
US-09-107-532A-2112/c
; Sequence 2112, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; CORRESPONDENCE ADDRESS:
; NUMBER OF SEQUENCES: 7310
; ADDRESSSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD/ROM ISO9660
```

```
;
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 2112:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 264 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (B) LOCATION 1...264
; SEQUENCE DESCRIPTION: SEQ ID NO: 2112:
US-09-107-532A-2112

Query Match 2.8%; Score 35.6; DB 4; Length 264;
Best Local Similarity 60.2%; Pred. No. 0.29;
Matches 59; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 1044 TTCCATAACTGCTGTGGGATTGCCTATAGTTCCAGTTGTAGCAGTGATTGTTTCCATGA 1103
Db 123 TTTTAGACTGTAGCGGAGTCTGCTTTTCTGTTGTTTAGAGATTGTTCCAAAAGT 64

QY 1104 TAAATGAACGGCTCAAGATCTTCCATCATTTTAG 1141
Db 63 GGAAATGATCTCTCACATTTCTTTTCACTGATTTCAG 26

RESULT 15
US-09-107-532A-3027
; Sequence 3027, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 3027:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 600 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (B) LOCATION 1...600
; SEQUENCE DESCRIPTION: SEQ ID NO: 3027:
US-09-107-532A-3027

Query Match 2.8%; Score 35.6; DB 4; Length 600;
Best Local Similarity 60.2%; Pred. No. 0.49;
Matches 59; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 1044 TTCCATAACTGCTGTGGGATTGCCTATAGTTCCAGTTGTAGCAGTGATTGTTTCCATGA 1103
Db 234 TTTTAGACTGTAGCGGAGTCTGCTTTTCTGTTGTTTAGAGATTGTTCCAAAAGT 293

QY 1104 TAAATGAACGGCTCAAGATCTTCCATCATTTTAG 1141
Db 294 GGAAATGATCTCTCACATTTCTTTTCACTGATTTCAG 331

Search completed: November 1, 2004, 22:40:49
Job time : 125.284 secs
```

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 1, 2004, 06:20:04 ; Search time 669.694 Seconds  
(without alignments)  
9900.026 Million cell updates/sec

Title: US-09-913-767-4  
Perfect score: 1293  
Sequence: 1 ggaagctctcttagtggtggt.....tatgatcaaacatatattcc 1293

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3413475 seqs, 253800928 residues  
Total number of hits satisfying chosen parameters: 6826950

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications\_NA:\*

1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*

2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*

3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*

4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*

5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*

6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*

7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*

8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*

9: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*

10: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*

11: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*

12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*

13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*

14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*

15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*

16: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq:\*

17: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq:\*

18: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*

19: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq:\*

20: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*

21: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1152	89.1	3387	9	US-09-938-842A-2212
2	1152	89.1	3387	11	US-09-938-842A-2212
3	369.4	28.6	1294	16	US-10-424-599-2524
4	354.4	27.4	592	9	US-09-770-152-44
5	307.6	23.8	2214	16	US-10-424-599-102519
c 6	295	22.8	1513	17	US-10-437-963-82061
c 7	269	20.8	2381	17	US-10-437-963-43860
8	257.6	19.9	997	16	US-10-424-599-75161
9	246.6	19.1	83698	17	US-10-416-898-9
10	162.2	12.5	792	16	US-10-424-599-9213
11	155.4	12.0	1386	16	US-10-425-114-12485
12	153.8	11.9	1779	16	US-10-425-114-28473

13	153	11.8	728	17	US-10-767-701-9000
14	152.4	11.8	522	16	US-10-424-599-114931
15	140.6	10.9	2605	17	US-10-437-963-83900
16	129.2	10.0	747	17	US-10-767-701-6806
c 17	128	9.9	386	9	US-09-770-791-85
18	122.6	9.5	1146	16	US-10-425-114-10516
19	122.6	9.5	1233	16	US-10-424-599-30274
20	121.4	9.4	671	17	US-10-767-701-5986
21	116	9.0	1432	17	US-10-437-963-1897
22	115.2	8.9	332	11	US-09-732-627A-1196
23	100	7.7	1333	16	US-10-425-114-10574
24	99.2	7.7	1373	16	US-10-424-599-31364
c 25	97	7.5	327	16	US-10-424-599-121152
26	96.6	7.5	650	17	US-10-767-701-4045
27	93.2	7.2	994	16	US-10-424-599-73501
28	89.6	6.9	1355	16	US-10-424-599-31363
29	72.6	5.6	1173	17	US-10-437-963-57539
30	70	5.4	716	16	US-10-424-599-61310
31	67.8	5.2	774	9	US-09-770-445-895
c 32	63	4.9	308	16	US-10-424-599-121805
33	61.8	4.8	2928	17	US-10-437-963-33800
34	60.2	4.7	1177	17	US-10-437-963-91425
35	58.2	4.5	789	17	US-10-437-963-72452
36	54	4.2	1417	16	US-10-424-599-97043
c 37	48.6	3.8	383	17	US-10-021-323-2053
38	47.4	3.7	1296	16	US-10-425-114-13687
39	46.8	3.6	382	17	US-10-021-323-592
c 40	46	3.6	2613	15	US-10-171-319-15
41	45.6	3.5	174424	9	US-09-967-788A-314
42	45.6	3.5	174424	10	US-09-960-706-969
c 43	43.6	3.4	2613	15	US-10-171-319-18
c 44	42	3.2	465	13	US-10-027-632-96997
c 45	42	3.2	465	13	US-10-027-632-306418

ALIGNMENTS

RESULT 1

US-09-938-842A-2212

; Sequence 2212, Application US/09938842A

; Patent No. US20020160378A1

; GENERAL INFORMATION:

; APPLICANT: Harper, Jeff

; APPLICANT: Kreps, Joel

; APPLICANT: Wang, Xun

; APPLICANT: Zhu, Tong

; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING

; FILE REFERENCE: SCRIPI300-3

; CURRENT APPLICATION NUMBER: US/09/938,842A

; CURRENT FILING DATE: 2001-08-24

; PRIOR APPLICATION NUMBER: US 60/227,866

; PRIOR FILING DATE: 2000-08-24

; PRIOR APPLICATION NUMBER: US 60/264,647

; PRIOR FILING DATE: 2001-01-16

; PRIOR APPLICATION NUMBER: US 60/300,111

; PRIOR FILING DATE: 2001-06-22

; NUMBER OF SEQ ID NOS: 5379

; SEQ ID NO 2212

; LENGTH: 3387

; TYPE: DNA

; ORGANISM: Arabidopsis thaliana

US-09-938-842A-2212

Query Match 89.1%; Score 1152; DB 9; Length 3387;  
Best Local Similarity 96.4%; Pred. No. 0;  
Matches 1207; Conservative 0; Mismatches 0; Indels 45; Gaps 1;

QY 22 ATGGAATAACTCAAGTAATCTATGTCATGCGTAAGCAAGATGCATCTCGTAGTGTAGAT 81  
|||||

DB 1 ATGGAATAACTCAAGTAATCTATGTCATGCGTAAGCAAGATGCATCTCGTAGTGTAGAT 60  
|||||

```
Qy 82 TACTTGATTTCTTTGCGTAACTCTGTTGTTTGTATGTTTTCAGATCATATAACAACATAGAA 141
Db 61 TACTTTGATTTCTTTGCGTAACTCTGTTGTTTGTATGTTTTCAGATCATATAACAACATAGAA 120
Qy 142 GCAAACTTAAACAGGTTCAGAGGAAATGAATACCAACCATGGAATCGAATCTTCTCGCGTA 201
Db 121 GCAAACTTAAACAGGTTCAGAGGAAATGAATACCAACCATGGAATCGAATCTTCTCGCGTA 180
Qy 202 CCTCAATCGAAGAACTATAAGAAATGGCTTCGTAATTTCCATTTACGTTGTTCTTTGTCCTT 261
Db 181 CCTCAATCGAAGAACTATAAGAAATGGCTTCGTAATTTCCATTTACGT----- 227
Qy 262 GCTTGCCCAAGCACATTCTACAAATTTGGCGAGATTTTACTATGAATAATGGTGGGAAGGT 321
Db 228 -----AGTTTACTATGAATAATGGTGGGAAGGT 255
Qy 322 ACATGGATGGGAACACATTGTCCAACTAATCGGCTTCCCTGTTCTGTTTCTCTCCGCTTC 381
Db 256 ACATGGATGGGAACACATTGTCCAACTAATCGGCTTCCCTGTTCTGTTTCTCTCCGCTTC 315
Qy 382 TTTTCCCAAAACAAAATCCAAACCAAGAGAGATTTTCAGAAAGTTCTCTTCCTTC 441
Db 316 TTTTCCCAAAACAAAATCCAAACCAAGAGAGATTTTCAGAAAGTTCTCTTCCTTC 375
Qy 442 ACCATTCTTGATCAGTTTACATCGTTACTGGACTATATAGTGTCTGCTAACTTATATG 501
Db 376 ACCATTCTTGATCAGTTTACATCGTTACTGGACTATATAGTGTCTGCTAACTTATATG 435
Qy 502 TCCTCTGTTGTTTACTATCTTACCAGTTTCTACTTTCTCCCTCATCTTGGCCTCAAA 561
Db 436 TCCTCTGTTGTTTACTATCTTACCAGTTTCTACTTTCTCCCTCATCTTGGCCTCAAA 495
Qy 562 TTGGCCTTCACTGCTTCTTCTCATATATTTCTAACTCGCAAGATTCACACCTTTCAAT 621
Db 496 TTGGCCTTCACTGCTTCTTCTCATATATTTCTAACTCGCAAGATTCACACCTTTCAAT 555
Qy 622 GTGAATTTCTGTTTCTCTTACTATTTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC 681
Db 556 GTGAATTTCTGTTTCTCTTACTATTTCTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCT 615
Qy 682 GAAACACAGCAAAAGTGTCTAGAGTAAATATGTGATAGGGAATAATATGATACCATTTGGT 741
Db 616 GAAACACAGCAAAAGTGTCTAGAGTAAATATGTGATAGGGAATAATATGATACCATTTGGT 675
Qy 742 GCTTCGTGCGGATTTGGATTTGCTGATFCCCTGGTACAACTGATCCTCAGGAAGTTTAA 801
Db 676 GCTTCGTGCGGATTTGGATTTGCTGATFCCCTGGTACAACTGATCCTCAGGAAGTTTAA 735
Qy 802 AAGAGCAAAACATTTCTCAACGGTCACTGACTTGGTCCCTTACCCTTCTAGTTTGCACG 861
Db 736 AAGAGCAAAACATTTCTCAACGGTCACTGACTTGGTCCCTTACCCTTCTAGTTTGCACG 795
Qy 862 TGTGTGGTTCTCATAGGACTTTTTCGAAGCGGGAGTGGAAAACTTTTAAACAGTGAGATG 921
Db 796 TGTGTGGTTCTCATAGGACTTTTTCGAAGCGGGAGTGGAAAACTTTTAAACAGTGAGATG 855
Qy 922 GAAACATCAAACTGGGGAAAGTGCCATACGTTATGACTTTTGGCTCGATAGCTATTTCC 981
Db 856 GAAACATCAAACTGGGGAAAGTGCCATACGTTATGACTTTTGGCTCGATAGCTATTTCC 915
Qy 982 TGGCAAGTCTACACCATTTGGGCTGCTGGGACTGATCTTTTGGTCTGATCTTCTGTTCTCC 1041
Db 916 TGGCAAGTCTACACCATTTGGGCTGCTGGGACTGATCTTTTGGTCTGATCTTCTGTTCTCC 975
Qy 1042 AATTCCATAAAGTGTGGGATTTGCTTATAGTTTCCAGTTGTAGCAGTGATTTGTTTCCAT 1101
Db 976 AATTCCATAAAGTGTGGGATTTGCTTATAGTTTCCAGTTGTAGCAGTGATTTGTTTCCAT 1035
Qy 1102 GATAAAATGAACGGCTCAAAAGATCTTCTCCATCATTTTATAGCTATCTGGGGATTCATTTCA 1161
Db 1036 GATAAAATGAACGGCTCAAAAGATCTTCTCCATCATTTTATAGCTATCTGGGGATTCATTTCA 1095
Qy 1162 TTTGTCATCAGCACTACCTCGACGAAAGATTTGAAGACTTAGCCACACAAAGTCTCTGTA 1221
```

```
Db 1096 TTTGTCATCAGCACTACCTCGAGAAAGAGTTGAGACTAGCCACACAAAGTCTCTGTA 1155
Qy 1222 GGAGATCCTCATCTACTACCTGCTGAGGAAGGTTCACACAAACATACATAGTG 1273
Db 1156 GGAGATCCTCATCTACTACCTGCTGAGGAAGGTTCACACAAACATACATAGTG 1207

RESULT 2
US-09-938-842A-2212
; Sequence 2212, Application US/09938842A
; Publication No. US20040009476A9
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 2212
; LENGTH: 3387
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-2212

Query Match 89.1%; Score 1152; DB 11; Length 3387;
Best Local Similarity 96.4%; Pred. No. 0;
Matches 1207; Conservative 0; Mismatches 0; Indels 45; Gaps 1;

Qy 22 ATGGAATAAATCAAGTAAATCTATGTCAATGTTGAAGCAAGATGTCATCTCGTAGTGTAGAT 81
Db 1 ATGGAATAAATCAAGTAAATCTATGTCAATGTTGAAGCAAGATGTCATCTCGTAGTGTAGAT 60
Qy 82 TACTTGATTTCTTTTCGCTAACCTGTTGTTTTCAGATGTTTTCAGATCATACACATAGAA 141
Db 61 TACTTGATTTCTTTTCGCTAACCTGTTGTTTTCAGATGTTTTCAGATCATACACATAGAA 120
Qy 142 GCAAACTTAAACAGGTTCAGGAGGAAATGAATACCAACCATGGAATTCGAATCTTCGTCGTA 201
Db 121 GCAAACTTAAACAGGTTCAGGAGGAAATGAATACCAACCATGGAATCTTCGTCGTA 180
Qy 202 CCTCAATCGAAGAACTATAAGAAATGGCTTCGTAATTTCCATTTACGTTGTTCTTTGTCCTT 261
Db 181 CCTCAATCGAAGAACTATAAGAAATGGCTTCGTAATTTCCATTTACGT----- 227
Qy 262 GCTTGCCCAAGCACATTCTCAAAATTTGGGCGAGATTTTACTATGAATAATGGTGGGAAGGT 321
Db 228 -----AGTTTACTATGAATAATGGTGGGAAGGT 255
Qy 322 ACATGGATGGGAACACATTGTCCAACTAATCGGCTTCCCTGTTCTGTTTCTCTCCGCTTC 381
Db 256 ACATGGATGGGAACACATTGTCCAACTAATCGGCTTCCCTGTTCTGTTTCTCTCCGCTTC 315
Qy 382 TTTTCCCAAAACAAAATCCAAACCAAGAGAGATTTTCAGAAAGTTCTCTTCCTTC 441
Db 316 TTTTCCCAAAACAAAATCCAAACCAAGAGAGATTTTCAGAAAGTTCTCTTCCTTC 375
Qy 442 ACCATTCTTGATCAGTTTACATCGTTACTGGACTATATAGTGTCTGCTAACTTATATG 501
Db 376 ACCATTCTTGATCAGTTTACATCGTTACTGGACTATATAGTGTCTGCTAACTTATATG 435
Qy 502 TCCTCTGTTGTTTACTATCTTACCAGTTTCTACTTTCTCCCTCATCTTGGCCTCAAA 561
Db 436 TCCTCTGTTGTTTACTATCTTACCAGTTTCTACTTTCTCCCTCATCTTGGCCTCAAA 495
Qy 562 TTGGCCTTCACTGCTTCTTCTCATATATTTCTAACTCGCAAGATTCACACCTTTCAAT 621
Db 496 TTGGCCTTCACTGCTTCTTCTCATATATTTCTAACTCGCAAGATTCACACCTTTCAAT 555
Qy 622 GTGAATTTCTGTTTCTCTTACTATTTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC 681
Db 556 GTGAATTTCTGTTTCTCTTACTATTTCTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCT 615
Qy 682 GAAACACAGCAAAAGTGTCTAGAGTAAATATGTGATAGGGAATAATATGATACCATTTGGT 741
Db 616 GAAACACAGCAAAAGTGTCTAGAGTAAATATGTGATAGGGAATAATATGATACCATTTGGT 675
Qy 742 GCTTCGTGCGGATTTGGATTTGCTGATFCCCTGGTACAACTGATCCTCAGGAAGTTTAA 801
Db 676 GCTTCGTGCGGATTTGGATTTGCTGATFCCCTGGTACAACTGATCCTCAGGAAGTTTAA 735
Qy 802 AAGAGCAAAACATTTCTCAACGGTCACTGACTTGGTCCCTTACCCTTCTAGTTTGCACG 861
Db 736 AAGAGCAAAACATTTCTCAACGGTCACTGACTTGGTCCCTTACCCTTCTAGTTTGCACG 795
Qy 862 TGTGTGGTTCTCATAGGACTTTTTCGAAGCGGGAGTGGAAAACTTTTAAACAGTGAGATG 921
Db 796 TGTGTGGTTCTCATAGGACTTTTTCGAAGCGGGAGTGGAAAACTTTTAAACAGTGAGATG 855
Qy 922 GAAACATCAAACTGGGGAAAGTGCCATACGTTATGACTTTTGGCTCGATAGCTATTTCC 981
Db 856 GAAACATCAAACTGGGGAAAGTGCCATACGTTATGACTTTTGGCTCGATAGCTATTTCC 915
Qy 982 TGGCAAGTCTACACCATTTGGGCTGCTGGGACTGATCTTTTGGTCTGATCTTCTGTTCTCC 1041
Db 916 TGGCAAGTCTACACCATTTGGGCTGCTGGGACTGATCTTTTGGTCTGATCTTCTGTTCTCC 975
Qy 1042 AATTCCATAAAGTGTGGGATTTGCTTATAGTTTCCAGTTGTAGCAGTGATTTGTTTCCAT 1101
Db 976 AATTCCATAAAGTGTGGGATTTGCTTATAGTTTCCAGTTGTAGCAGTGATTTGTTTCCAT 1035
Qy 1102 GATAAAATGAACGGCTCAAAAGATCTTCTCCATCATTTTATAGCTATCTGGGGATTCATTTCA 1161
Db 1036 GATAAAATGAACGGCTCAAAAGATCTTCTCCATCATTTTATAGCTATCTGGGGATTCATTTCA 1095
Qy 1162 TTTGTCATCAGCACTACCTCGACGAAAGATTTGAAGACTTAGCCACACAAAGTCTCTGTA 1221
```

Db 436 TCCTCTGTTGGTTTACTATACCTACAGTTTCTTCTCTCCCTCATCTTTGGCCTCACAA 495  
Qy 562 TTGGCCCTTCACTGCCCTTCTTCTCATATTTTCTAAACTCGCAGAAAGTTTCAACCTTTCAAT 621  
Db 496 TTGGCCCTTCACTGCCCTTCTTCTCATATTTTCTAAACTCGCAGAAAGTTTCAACCTTTCAAT 555  
Qy 622 GTGAATCTCTGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 681  
Db 556 GTGAATCTCTGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 615  
Qy 682 GAAACACACGAAAGTGTCTACAGTAAATATGTATAGGGAATATATGTACCATTTGGT 741  
Db 616 GAAACACACGAAAGTGTCTACAGTAAATATGTATAGGGAATATATGTACCATTTGGT 675  
Qy 742 GCTTCTGCTGGGATTGGATTGCTGCTATCCCTGGTACAACTGATCCTCAGGAAGTTTAA 801  
Db 676 GCTTCTGCTGGGATTGGATTGCTGCTATCCCTGGTACAACTGATCCTCAGGAAGTTTAA 735  
Qy 802 AGAAGCAAACTTCTCAAGGTCATCTGATCTGGTTCGCTTACCAATCTCTAGTTGCAAGC 861  
Db 736 AGAAGCAAACTTCTCAAGGTCATCTGATCTGGTTCGCTTACCAATCTCTAGTTGCAAGC 795  
Qy 862 TGTGTCGTTCTCATAGGACTTTTCGCAAGGGGGAGTGGAAACTTTTAACTTAACTAGATG 921  
Db 796 TGTGTCGTTCTCATAGGACTTTTCGCAAGGGGGAGTGGAAACTTTTAACTTAACTAGATG 855  
Qy 922 GAAACTACAAACTCGGGGAAAGTGCATACGTTATGACTTTTGGCCCTCGATAGCTATTTCC 981  
Db 856 GAAACTACAAACTCGGGGAAAGTGCATACGTTATGACTTTTGGCCCTCGATAGCTATTTCC 915  
Qy 982 TGGCAAGTCTACCACTATGGCGTGGGACTGATCTTTGAGTCACTTTCTGTTCTCC 1041  
Db 916 TGGCAAGTCTACCACTATGGCGTGGGACTGATCTTTGAGTCACTTTCTGTTCTCC 975  
Qy 1042 AATTCATTAAGTCTGCTGGGATGCTATAGTTTCCAGTTTGGAGTGGTTTCCAT 1101  
Db 976 AATTCATTAAGTCTGCTGGGATGCTATAGTTTCCAGTTTGGAGTGGTTTCCAT 1035  
Qy 1102 GATAAATGAACGCTCAAGATCTTCTCATCATTTTATAGCTATCTGGGGATTTCATTTCA 1161  
Db 1036 GATAAATGAACGCTCAAGATCTTCTCATCATTTTATAGCTATCTGGGGATTTCATTTCA 1095  
Qy 1162 TTTGCTATCAGCACTTACCTCGACGAAAGAGTTGAAGACTAGCCACAAAGTCTGTA 1221  
Db 1096 TTTGCTATCAGCACTTACCTCGACGAAAGAGTTGAAGACTAGCCACAAAGTCTGTA 1155  
Qy 1222 CGAGATCCTCATCTACTCTCTGAGGAAGTGCACAAACATACATAGT 1273  
Db 1156 GGAGATCCTCATCTACTCTCTGAGGAAGTGCACAAACATACATAGT 1207

RESULT 3  
US-10-424-599-2524  
; Sequence 2524, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Roesa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 2524  
; LENGTH: 1294  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_102285C.1  
US-10-424-599-2524

Query Match 28.6%; Score 369.4; DB 16; Length 1294;  
Best Local Similarity 60.1%; Pred. No. 1.2e-98;  
Matches 644; Conservative 0; Mismatches 406; Indels 21; Gaps 1;  
Qy 211 AAGAACTATAAGAAATGGCTTCGTATTTCCATTATTCAGTGTCTTTGTCTCTTGTCTGCCAA 270  
Db 119 AAAAGATATTACCGTTGGCTCAGAATATCCATCCATTTCTTCACTTGTGTGTGCGGA 178  
Qy 271 GCACITTCACAAATTTGGGCGAGGTTTACTATGAAATGGTGGGAGAGTACATCGATG 330  
Db 179 TCAGCAGCAATATCTCTTGGAAAGATTGTACTATGAAAGGGTGGAAAAAGCAAGTGGATG 238  
Qy 331 GGAAACATTTGTCCAACTAATCGCTTCCCTGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 390  
Db 239 GGAAACATTTGTCAACTTCTGCTGTTTCCCTATTTAGCTTCTTCTTCTTCTTCTTCTTCTTCT 298  
Qy 391 ACCAAATATCCAAACCAACAGAGAGATTTTCAAGAAAGTTTCTTCTTCTTCTTCTTCTTCTTCT 450  
Db 299 TCCAAAAATCTCACCACAAATAGTAGCATTTTCAAAACAATCATCTCTTCAATACTA 358  
Qy 451 GGATCAGTTTACATCGTTACTGGACTATTTAGTGTCTGCTAACTCTTATATATGTCCTCTGTT 510  
Db 359 GCAITTTATCTATGTCTCAATTTGGCCTACTTTTGGCATTTAGATTGTATTTATATTACAGTT 418  
Qy 511 GGTTTACTATATCTTACCAGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 570  
Db 419 GGACTATGGGACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 478  
Qy 571 ACTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 630  
Db 479 AATGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 538  
Qy 631 CTGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 680  
Db 539 CTAGTCT 598  
Qy 681 -----GAAAAACACAGCAAAAGTGTCTAGAGTAAAAATATGTATAGGGATAATA 729  
Db 599 GATGATGATGATTAGATTCCACAAAATCTCAGAAAAAGATATGTATAGGATTCATA 658  
Qy 730 TGTACATTTGTTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 789  
Db 659 TGCAGATTGGTGATCTGCGGGTATGGACTATGGCTTTCCTTCCCTCACACAGCTTGTGTTCT 718  
Qy 790 AGGAAGTTTAAAGAACAAACATTTCAACGGTCTGACTTGGTCTGCTTACCAATCT 849  
Db 719 AAGAAAGTCAATAAAAGAGAAACATTCAAGTGGTCTTGGATATGATATATACACTTCC 778  
Qy 850 CTAGTTGCAAGCTGTGTGTTCTCATAGGACTTTTTCGCAAGCGGGAGTGGAAAACTTTA 909  
Db 779 CTTGTGGTACCTTGTCTACCTTAGTGGACCTTTTGTAGTGGAGAGTGGAGTGGTTG 838  
Qy 910 ACAAGTGAAGTGAAGAACTACAACTGGGGAAGTGCATACGTTTATGACTTTTGGCCTCG 969  
Db 839 AAGAAATGAAGAGTATGAGTTGGGAAGGCTTCAATTTTCTGCTGAACCTCACTTTC 898  
Qy 970 ATAGCTATTTCTGSCAAAGTCTACACATTTGGCGTCTGGGACTGATCTTTGAGTCACT 1029  
Db 899 ACAGCATACTTTGGCAAGTCTTCCACTTGGTGTGTTTGGGCGATATTTAGTGAAGTTC 958  
Qy 1030 TCTGTGTTTCTCAATTTCCATAACTCTCTGTTGGATTGCTTATAGTTTCCAGTTCTAGCAGTG 1089  
Db 959 TCCCTCTTCTTAATGCCATAAGCGCTTTGGGAGTGCCTATTTGTTCCAAATGCTGGCAGTG 1018  
Qy 1090 ATTGTTTCCATGATAAAATGAACGCTCAAGATCTTCTCATCATTTTAGCTATCTGG 1149  
Db 1019 TTGTTCTTTCATGACAAAATGGATGGCATTAAGGGTATCTCTATGGTGTAGCTATTTGG 1078  
Qy 1150 GGATTCATTTCTATTTGCTCTATCAGCACTACCTCGACGAAAGAGAGTTGAGNACTAGCCAC 1209  
Db 1079 GGCATTGTATCATATGTGTATCAACAGTACTTGGATGATACAAAGTCCGAAAAACAAAC 1138

Qy 1210 ACAAGTCTGTAGGATCCTCATCTACTTACCTGCTGAGGAAGGTACACA 1260

Db 1139 ACTACTAGTCATGTTCCAAAAGTCTCTTCACTATAGAAAGGTTTCATAGA 1189

```

RESULT 4
US-09-770-152-44
; Sequence 44, Application US/09770152
; Publication No. US20020040489A1
; GENERAL INFORMATION:
; APPLICANT: Gorlach, Jorn
; APPLICANT: An, Yong-Qiang
; APPLICANT: Hamilton, Carol M.
; APPLICANT: Price, Jennifer L.
; APPLICANT: Raines, Tracy M.
; APPLICANT: Yu, Yang
; APPLICANT: Rameaka, Joshua G.
; APPLICANT: Page, Amy
; APPLICANT: Matthew, Abraham V.
; APPLICANT: Ledford, Brooke L.
; APPLICANT: Woessner, Jeffrey P.
; APPLICANT: Haas, William David
; APPLICANT: Garcia, Carlos A.
; APPLICANT: Kricker, Maja
; APPLICANT: Slader, Ted
; APPLICANT: Davis, Keith R.
; APPLICANT: Allen, Keith
; APPLICANT: Hoffman, Neil
; APPLICANT: Hufban, Patrick
; TITLE OF INVENTION: Expressed Sequences
; FILE REFERENCE: 2025US (PARA-014PRV)
; CURRENT APPLICATION NUMBER: US/09/770-152-44
; CURRENT FILING DATE: 2003-01-26
; PRIOR APPLICATION NUMBER: 60/178,500
; PRIOR FILING DATE: 2000-01-27
; NUMBER OF SEQ ID NOS: 999
; SOFTWARE: FastSeq for Windows Version 1.0
; SEQ ID NO 44
; LENGTH: 592
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-770-152-44

```

		Query Match	27.4%; Score 354.4; DB 9; Length 592;	
		Best Local Similarity	81.8%; Pred. No. 2.2e-94;	
		Matches	Conservative 0; Mismatches 91; Indels 0; Gaps 0;	
		Matches	409; Conservative 0; Mismatches 91; Indels 0; Gaps 0;	
Qy	180	GGAATCGAATCTTCGTCGGTACCTCAATCGAAGAACTATAAGAAATGGCTTCGTATTTC	239	
Db	93	GGAATCGAATCATTTTCAGTAGCTCAAAACGAGAATGTAAAGAGTGCCCTCCGTGCTC	152	
Qy	240	CATTTACGTGTCTTTTGTCTTGTGCGCAAGACACTTTCTACAATTTTTGGGCAGAGTTTA	299	
Db	153	CATATACGCAATCTTTGTGCATCTTCTGCCAACCACTTGCTACAGTCTTGCGGTAGACTGA	212	
Qy	300	CTATGAAAATCGTGGGAAGAGTACATGGATCGGAACACTTGTCCAATAATCGGCTTCCC	359	
Db	213	CTATGAAAATGGAGGAAAAAGCACATATGTGGTAAACACTTCTTCAAACCTCATTTGGCTTCCC	272	
Qy	360	TGTTCTGTCTTCTCTCCGCTCTTTTCCCAACCAGAAAATCCCAAAACCAACAGAACAGA	419	
Db	273	TGTAAGTATCTCTTCGCTTCTTTTCTCGAATCAGGCAACCCCAAAATCAACAGATACAAA	332	
Qy	420	TTTCAGAAAGTCTCTTCCTTCCACATCTTTGGGATCAGTTTACATCGTTACTGACATTT	479	
Db	333	TTTCAGTACAGTCCCTTCCITCACACOCCTTGCATCGGTTTACTTGTGCATCGACTGCT	392	
Qy	480	AGTGTCTGCTAACTCTTATATGTCTCTGTGGTGTACTACTACTTACCAGTTTCTACTTTT	539	
Db	393	AGTGTCTGCTTATGCTTATTTTGTCTGCAAGTTGGGTGCTTTTACTTACCACGCTCTACTTTT	452	
Qy	540	CTCCCTCAICTTGGCCTCACAAATGGCCCTTCACGTCCCTTCTCTCATATTTTCTAAACCTC	599	

Db 453 CTCCTCACTCTGGCCTCACAGTTGGCCCTTCTCACTGCGCTTTCTCTATATTTCTTAACTC 512

Qy 600 GCAGAAAGTTCCACACCTTTTCATTTGGAATCTCTGTGTTCTCCTTACTATTTCTCTGCGCT 659

Db 513 GCAAAAGTTCACTCTTTGTAGTCAATCTTTGTTCTCCTTACGGTTTCTCTGCGCT 572

Qy 660 CCTCGTGGTCAACACTGATT 679

Db 573 CCTCGTGGTCAACACTGATT 592

RESULT 5

US-10-424-599-102519

; Sequence 102519, Application US/10424599

; Publication No. US20040031072A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa Thomas J

; APPLICANT: Kovalic David K

; APPLICANT: Zhou Yihua

; APPLICANT: Cao Yongwei

; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with Soybean

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53223)B

; CURRENT APPLICATION NUMBER: US/10/424,599

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 285684

; SEQ ID NO 102519

; LENGTH: 2214

; TYPE: DNA

; ORGANISM: Glycine max

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_63592C.1

US-10-424-599-102519

Query Match	23.8%;	Score 307.6;	DB 16;	Length 2214;
Best Local Similarity	58.5%;	Pred. No. 4e-80;		
Matches 580;	Conservative 0;	Mismatches 399;	Indels 13;	Gaps 2;
QY	204	TCAATCAAGAACATATAAGAAATGGCTTCGTATTTCATTTCCATTTACGTGTTCTTTTGTCCCTTGC	263	
DB	389	TAAATATAAGCGATGGCAATGGTGGTCTCTGTGACACTCAGCATAGCCTTTCTTATAGT	448	
QY	264	TGCGCAAGCACTTTCTCAAAATTTGGGCGAGATTACTATGAAAAATGGTGGGAAGAGTAC	323	
DB	449	TGGGCAATCTGCTGCTGTATCTCTTGGAAAGATTTTATTATGATCAGGGTGGAAATAGTAA	508	
QY	324	ATGATATGGGAACACTGTGTCAACTAAATCGGCTCCCTTGTTCTGTTTCTCTTCCCGCTTCTT	383	
DB	509	ATGATATGGCTACTCTAGTTTCAAACCTGCTGCCCTCCGATCTTG-----TTTCAT	556	
QY	384	TTCCCAACCAAAAATCCCAACCAAGACAGATTTTCAGAAAGTGTCTCTCTCTTCAC	443	
DB	557	TCCATTTATTACAATTCCTTCACCTCCAGAGGCTTCAACTTCTGCTTCACCTCCCATCAA	616	
QY	444	CATTCTTGGATCAGTTTACATCGTTAC-TGGACTATTAGTGTCTGCTAACTCTTATATGT	502	
DB	617	AAATTAATCTTTTGTATATTTTGGTCTTTGGAGTCTTAATTGCCTGCACAAATATGATGT	676	
QY	503	CCTCTGTTGGTTTACTACTTACACAGTTTCTACTTTCTCCCTCATCTTGGCCCTCACAAT	562	
DB	677	ACTCCACTGGACTCTTATACCTCTCGGCTCTACCTATTCCGTGATTTGTGCATCACAGT	736	
QY	563	TGGGCTTCACATGCTTCTTCTCATATTTTCTAAACTGCAGAAAGTTTCACACCTTTCAATG	622	
DB	737	TAGCTTTTAAATGCAAGTTTCTCTCATATTTTATCAAATTCCTCAAAGTTTCACTGSCCTTGATTA	796	
QY	623	TGAATCTCTGTTTCTCCTTACTATTTTCTCTGCGCTCCTCGTGGTCAACACTGATTCGG	682	
DB	797	TAAACTCTACAGTGGTTCTCACTTTATCTGCTGCACCTCTTGTGTTTAAACGAAACACAG	856	
QY	683	AAAAACACAGCAAAAAGTGCTAGAGTAAATATGTGATAGGGATTAATATGTTACCAATTTGGTG	742	

Db 857 ATGAACCATCTGGTTCTCCAAAGGAAAGTACATTAATGGTTTCCATGTACCCCTTGGAG 916  
Qy 743 CTTCTGCTGGGATGGATGCTCTATCCCTGTGTACAACTGATCCTCAGGAAGGTTTTAA 802  
Db 917 CTTCTGAGTGTAATCTCTTTTGGCTTTCCCTCATGAGCTGACCTTTGAGAGGTTCTGA 976  
Qy 803 AGAAGCAAAATTTCTCAAGGTCACCTGATGCTGTGCTGTACCAATCTCTAGTTGCAAGCT 862  
Db 977 AGAAGGAAACATTTCTGTGTTTGGAAATGCAAAATCTACACATCAATTCGTTGCCTCTG 1036  
Qy 863 GTGTGTTCTCATAGGACTTTTCGCAAGCGGAGTGGAAACCTTTAAACAAGTGAGATGS 922  
Db 1037 GTGCTCTGTATAGGCTATTTGCAAGTGGGAATGGCGTACTTTGCAATGGAGAATGG 1096  
Qy 923 AAAAATAACAACTGGGAAAGTCCCATAGCTTATGACTTTTGGCTCGATAGCTATTTCTCT 982  
Db 1097 AGGGTTTTCAGAAAGGATATGTTGCTTATGTTATGACTTTGGTTGGAATTCGAATAGCT 1156  
Qy 983 GGCAGTCTACCAATTTGGGCTGGGACTGATCTTTGAGTCATCTCTGTGTTCTCCA 1042  
Db 1157 GGCAGGTATGCTCTGTGTTGGTGTGTTGGCTTGATCTTCTAGTGTCTTCTCTACTCCA 1216  
Qy 1043 ATTCCATACTGCTGTGGGATTCCTCTATAGTTCAGTTCAGTGTGAGTGTGTTTCCATG 1102  
Db 1217 ATGTTATAGCAAGTTCTTTAGCCGTAATCTCTATGCTGTGTTATAGTGTGTTTCTATG 1276  
Qy 1103 ATAAATGAACGCGTCAAAAGATCTTCTCCATCATTTTATGCTATCTGGGATTCATTTTCA 1162  
Db 1277 ATAAAGATGAATGGGTGAAGATAATTTCTATGCTTTTGGCTCTATGGGTTTGGCTCTT 1336  
Qy 1163 TTGCTATCAGCACTACCTCGAGCAAAAGAG 1194  
Db 1337 ATATTTATCAGAATTATCTTGATGATTCAAAG 1368

## RESULT 6

US-10-437-963-82061/c  
; Sequence 82061, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 82061  
; LENGTH: 1513  
; TYPE: DNA  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_81525C.1  
US-10-437-963-82061

Query Match 22.8%; Score 295; DB 17; Length 1513;  
Best Local Similarity 56.9%; Pred. No. 1.8e-76;  
Matches 563; Conservative 0; Mismatches 420; Indels 6; Gaps 1;  
Qy 211 AAGAACTATAAGAAATGGCTTCGTAATTTCCATTTACGTTGTTCTTTGTCTCTGCTTGCCTG 270  
Db 1270 AAGAGTTGCAATGGTGGCTGATGGTGGGAGTTAAACATGTTCTTCTCTCATTCGCGTCTAG 1211  
Qy 271 GCATTTCTCAATTTTGGGAGAGTTTACTATGAAAATGGTGGGAGAGTACATGGATG 330  
Db 1210 ACCGCATCTACATTTCTTGGGAGGTTCTACTCAATCAAGGGGCAATAGCAAGTGGATG 1151

Qy 331 GGAAACACTGTGTCAACTAATCGGCTTCCCTGTTCTGTTTCTCTTCGCTTCTTTTCCCAA 390  
Db 1150 TCCACATCGTCCAAACTGCTGCTTCCGATTTTGTTCATTTGCCCTATTTCTTTTCCAT 1091  
Qy 391 ACCAAAATCCCAACCAACAGAGAGATTTTCAGAAAGTTCTC-----TTCCTTCACC 444  
Db 1090 TCAAAGACATCTTCTACAAACAGTCACTAGTAGTCTCTGCCCTTACAAATTTCTATCCCC 1031  
Qy 445 ATTTCTGGATCAGTTTACATCGTTACTGACTATTAGTCTGCTAACTCTTATATGTCC 504  
Db 1030 AAAATTTACTGTATATATGTTGCTCTGGSCCTCATCATTTGCTGACAGCACTTGATGAT 971  
Qy 505 TCTGTGTTTACTACTTACCTTACAGTTTCTACTTTCTCCCTCATCTTTGGCCTCACAATTTG 564  
Db 970 TCTATGGCTACTATATCTTCGGGTCTCAACATATTCGCTCATCTGTGCTAGTCACTTT 911  
Qy 565 GCTTCACTGCTTCTTCTCATATTTTCTAACTCGCAGAGATTTCAACACTTTTCAATTTG 624  
Db 910 GCTTCAATGCTGCTCTCTCATATTTCTCAATGCTCAAAAATTCACCCCTCTGATTTTC 851  
Qy 625 AATTCTCTGTTTCTCTTACTATTTTCTCTGCTGCTGCTCAACACTGATTCGGA 684  
Db 850 AATTCCGTAGTCTCTTACGTTTCTGCTTCACTCTTGGAGTTGATGAAGATTTCTCAG 791  
Qy 685 AACACAGCAAAAGTGTCTAGAGTAAATATGTAGAGGATATATGTACCATTTGGTGCT 744  
Db 790 GGAATCTACTAGTATATCATATGGAAGTACATTTTGGGTTTCTGTTGACACTAGGGCA 731  
Qy 745 TCTGCTGGATGGATGCTGCTATCTCCCTGTTACAACTGATCTCTCAGGAAGTTTAAAG 804  
Db 730 TCAGCTCATACTCTGCTCATTTCTCTCCCTGATGCAAGTCAATTTGAGAAGGTTATTAAG 671  
Qy 805 AAGCAACATCTCTCAACGCTCACTGCTGCTTACCAATCTCTAGTTCGAGCTGT 864  
Db 670 AGGGAGACCTTCTCAGTTGTGTGAACATGCAATATATACAGCTCTCTGTCGCAACATTTG 611  
Qy 865 GTGGTTCTCATAGGACTTTTCCAAAGCGGAGTGGAAACCTTTAAACAGTGAAGTGA 924  
Db 610 GCTTCTCTGTTGGGTATTTGCAAGTGGTGAATGATGACTTTACAGGAGAGATGCAT 551  
Qy 925 AACTACAACTGGGAAAGTGCATACGTTATGACTTTTGGCTCGATAGCTATTTCTG 984  
Db 550 GCATTTCCAGTCTGGGAAGCTGTATATGTAATGACACTGCTGGACGCTATATCTTGG 491  
Qy 985 CAAGTCTACACCATTTGGGCTGCTGGGAGTGTCTTTGAGTCACTCTCTGTTCTTCCAT 1044  
Db 490 CAGGTAGCATCAGTTGGAGTGGTGGATTTGATCTTTTGTGTTGTCATCGCTGTTTCAAAT 431  
Qy 1045 TCCATAAATGCTGTGGGATTTGCTATAGTTCCAGTTGTAGCAGTGAATTTTTCATGAT 1104  
Db 430 GTGATAAGCACCTTAGCTTACCCATCATCTCTGTTTGTGCTGATTTTCTTTCATGAC 371  
Qy 1105 AAATGAACGGCTCAAGATCTTCTCCATCATTTTATGCTATCTGGGATTCATTTCAAT 1164  
Db 370 AAGATGGATGGAGTAAAGATTATAGCTATGCTGATGGCAATTTGGGATTTATGTCATAT 311  
Qy 1165 GTCTATCAGCACTACTCTCGCAAAAGAA 1193  
Db 310 GGCCCAACCAATTATATGTTGATGGCAAGAA 282

## RESULT 7

US-10-437-963-43860/c  
; Sequence 43860, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.



D	b		S	33	CAGAAATTCACAGCAATTCTGTTGTCTTCCCTTACCATAATCAGACTTCTCTG	952
Q	y			661	CTCGTGTCGAACACTGATTC---GGAAAAACAACAGCAAAGTGTCTAGAGTAAGAATATGTG	717
D	b			593	CTTGCAATCAACTCCGATTCGCAGCGAGGACTCAACCGGTCTTTCCAGAGAAAAGCATGTA	652
Q	y			718	ATAGGGATAAATATGTACCAITTTGGTCTTCTGCTGGGATTTGGATTGCTGTATCCCTGGTA	777
D	b			653	ATTGGATTCTTCTCACCAANTTTGGTGGCATCAGTACATTCTCTTTGTACCTCTCTCTTGTG	712
Q	y			778	CAACTGATCCTCAGGAGTTTTTAAAGAACCAACATTTCTCAACGGTCACTGACTTTGGTC	837
D	b			713	CAGCTTCTTTTCAGAAAGTTATAAGAGAGAAACCTTTTCTGCTGTTGGACATGCCAA	772
Q	y			838	GCTTTACCAATCTTAGTTTCCAAGCTGTGTGTTCTCATAGGACTTTTTTCGCAAGCGGGAG	897
D	b			773	TMTTACCACCTCTTCAITTTGCTACATGTGCTGTAGTGGTGTGTTGCAAGTGGAGAG	832
Q	y			898	TGGAAACTTTAACAAGTAGATGGA AAA CTACAA ACTGGG AA AG TG CC AT AC GT T A TG	957
D	b			833	TGGAAAAGTTTGAACAATAGATGAAGGGATATGATAAGGGAAGTGTGCCCGTTGTAATG	892
Q	y			958	ACTTTGGCTCGATAGCTATTCTCTGCAAGTCTACACCATTGCGTCTGGGACTGATC	1017
D	b			893	ACTGTACTATAGATTGCTGTGACATGGTATATATGCTCACTAGGGATGCTAGGGGGAGC	952
Q	y			1018	TTTGAGTCATCTTCTGTGTTCTCCAAITTCATAACTGCTGTGGG	1061
D	b			953	TTGGAGGAGTCTGCAITTTCTCGATGTGGAAGTACCTTGGGG	996
<b>RESULT 9</b>						
US-10-416-898-9						
; Sequence 9, Application US/10416898						
; Publication No. US20040172670A1						
; GENERAL INFORMATION:						
; APPLICANT: Yale University						
; APPLICANT: Walker, Elisabeth						
; APPLICANT: Dellaporta, Stephen						
; TITLE OF INVENTION: MAIZE YELLOW STRIPE1 AND RELATED GENES						
; FILE REFERENCE: 44574-5106-US						
; CURRENT APPLICATION NUMBER: US/10/416, 898						
; CURRENT FILING DATE: 2003-05-16						
; PRIOR APPLICATION NUMBER: PCT/US01/43101						
; PRIOR FILING DATE: 2001-11-16						
; PRIOR APPLICATION NUMBER: US 60/249,222						
; PRIOR FILING DATE: 2000-11-16						
; NUMBER OF SEQ ID NOS: 20						
; SOFTWARE: PatentIn version 3.1						
; SEQ ID NO 9						
; LENGTH: 83698						
; TYPE: DNA						
; ORGANISM: Zea mays						
US-10-416-898-9						
Query Match                19.1%; Score 246.6; DB 17; Length 83698;						
Best Local Similarity     53.0%; Pred. No. 4.9e-61;						
Matches      582; Conservative      0; Mismatches      504; Indels      13; Gaps          2;						
Q	y			99	TAACTGTGTTTTTGATGTTTTTCAGATNCATAACAA CATAGAGCA AAC CTAACAGGTCA	158
D	b			67100	TGAAC TTTTATGTCTGTTT TTGTGTGTAACAGATGAAGAAGCAATGTTCTGCTTA	67155
Q	y			159	GGAGGAATGAATACCA ACCAGTAAGAA TC GAATCTT CGTCGTACCTCAATCGAAGACTA	218
D	b			67160	AGGAAGAAGATGAAGGAAGACGAAGA ACTAGTGTCTCTACAACTCAATAA TA	67215
Q	y			219	TAAGAAA----TGGCTTCGATTTTCCCATTTTACGTGTTCTTTTGTCTGTGTCACCAAGCAC	274
D	b			67220	GAACCAATGGTGAATCTTGTTTTTCATAAGCATCTCTTTCTCATCTCTGTCTCAAGCCA	67270
Q	y			275	TTTCTCAAA TTTTGGG CAGAGT TTTACTATGAAAAT GGTTGGGAAGAGTACATGATGGGAA	334

D	b		67280	TTTCTGTTCTTCTTGTCGGTTGTATTAACAAGAGGTGGAAACAGTAAAAATGGATCTCTA	67333
Q	y		335	CACTTGTCCAACTAAATCGGCTTCCCCTGTTCTGTCTCTTCCGCCTCTTTTCCCAACCA	394
D	b		67340	CTCTTGTTCAAACTGGTGGCTTCCCATTTCTTTATCTCCTCTTTCTCTCGTTC-----	67393
Q	y		395	AAAATCCCAAACCAACAGAAGCAGATTTCAGAAAAGTTCTCTTCTCTCACCATTTCTGGAT	454
D	b		67394	--CTGCTTCAAAATCTTCTTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCAAGACTCTGGTTT	67450
Q	y		455	CAGTTTACATCGTTTACTGCACTATTAGTGTCTGCTAACCTTTATATATGTCTCTGTGGT	514
D	b		67451	GGATTATCTTTCTCTTGGTTTTGCTATTGGTTTAGATAAATTTTTTATAGTCTGTGGAC	67510
Q	y		515	TACTATACTTACCAGTTTCTACTTTCTCCCTCAATCTTGGCTCACAAATGGCCTTCACTG	574
D	b		67511	TTTTGTATCTCTCTGCTCAACTTATTCGATCTATGTGTTTCTCACAGTTAGCTTTCAATG	67570
Q	y		575	CCTTCTTCTCATATTTTCTAAACTCGCAGAAGTTCCACCTTTTCATTGTGAATTTCTCTGT	634
D	b		67571	GTGCTCTTCTATTATACATCAATTCACGAAAATCACTTGTGTGATTTTCTTCTCACTGT	67630
Q	y		635	TTCTCTTACTATTTTCTCTGCCCCCTCGTGGTCAACACTGATTCGGAACACACGAA	694
D	b		67631	TGTTTCTCTATCTCTGCTGTGTGGTTTCTCTTGACGATGANTCAAAATAGCCCATCAG	67690
Q	y		695	AAGTCTCTAGAGTAAAAATATGTATAGGATAAATATGTACCATTTGGTCTTCTGCTGGGA	754
D	b		67691	GAGATTTAAGTGGAGTTACTTGATTTGGGTGTTTTCTGTGCAGTTTTTGTCTTCTTATCT	67750
Q	y		755	TTGGATTGCTGCTATCCCTGGTGTACAACTGATCCTCAGGAAGGTTTTTAAAGAAGCAACAT	814
D	b		67751	ATTCTCTTCAGCTCTCTCTATGCACTTTTCTTCGAGAAGGTTCTCAAGAGTGAGACTC	67810
Q	y		815	TCTCAACGGTCACTGACTTGTGTCGCTTACCAATCTCTAGTTGCAAGCTGTGTGGTTCTCA	874
D	b		67811	TCTCTATGGTTCTCAGAGATGCAAAATCTATACGTGCTGTGGCTTCTTGTGTAGCGGTTA	67870
Q	y		875	TAGGACTTTTCGCAAGCGGGAGTCGMAAACTTTTAACAAGTGAGATGGAACATACAAAC	934
D	b		67871	TCGGATTGTTTCGAAGCGGGGAATGGAATGTTGTGTAGTGTGGAGATGGAAGATTTTCAGG	67930
Q	y		935	TGGGAAAAGTGCATPACGTTATGACTTTTGGCCTCGATAGCTATTTCTCGCAAGTCTACA	994
D	b		67931	AAGGTCAAGTCATTTATGTTTGTGACTTTTGGTCGGGCAGCGGTTTCGTGTCAATGGGTT	67990
Q	y		995	CCATTGGCGTCGTGGGACTGATCTTTTGAGTCATCTTCTGTGTCTCCAATCCATAACTG	1054
D	b		67991	GTGTAGGAGCGGTGCGCTTATATTCTCGTGTCTTTCGTGTTTTCAAAACCTTATTAGTA	68050
Q	y		1055	CTGTGGGATTGCCATAGTTCAGTTGTAGCAGTGTATTTTCCATGATAAAATGAACG	1114
D	b		68051	CGCTCTCACTCAATTTGTACGCCCTCTCGCGGCCAATTCGCGTGTTCATGACAAGTGA	68110
Q	y		1115	CGTCAAAGATCTTCTCCCATCATTTTATAGTCTATCTGGGATTCATTTTCTATCTCAGC	1174
D	b		68111	AGGTTAAGATGGTGGCGATGCCCATCGCCTTCACAGGATTCACGTTTTATATCTACCAGA	68170
Q	y		1175	ACTACCTCGACGAAAGAA	1193
D	b		68171	ACTATCTTTGATGACTTTGAA	68189

RESULT 10  
US-10-424-599-9213  
; Sequence 9213, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalich David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Sov Nucleic Acid









Db	2093	GGGTGTCGTCGCTGTTGCGCCAACTGACGGGCACGCTGTCGCTGCGCGCTGGTGCCGGTGT	2152
Qy	1082	TAGCAGTGAATTGTTTTCATGATAAAATGAACGCGTCAAGATCTTCTCCATCATTTTAG	1141
Db	2153	TCGCCGTCGCGCTGTTTCGGGGACAGGATGACCGGGATCAAGGCCGTATCCATGCTCATGG	2212
Qy	1142	CTATCTGGGGATTTCATTTGTTCTATCAGCACTACCTCGACG	1186
Db	2213	CCGTCTGGGGTTTCCTCTCGTACGCGTACCAGCAGTACATCGACG	2257

Search completed: November 1, 2004, 23:40:02  
Job time : 674.694 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 1, 2004, 05:45:29 ; Search time 4606.44 Seconds  
(without alignments)  
10228.408 Million cell updates/sec

Title: US-09-913-767-4

Perfect score: 1293  
Sequence: 1 ggaagctctcttagtggtggt.....tatgatcaaacatattccc 1293

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:  
1: gb\_est1:  
2: gb\_est2:  
3: gb\_hic:  
4: gb\_est3:  
5: gb\_est4:  
6: gb\_est5:  
7: gb\_est6:  
8: gb\_gss1:  
9: gb\_gss2:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	806	62.3	1340	3 CNS0A29J	BX827536 Arabidops
2	778	60.2	1431	3 CNS0A2KR	BX829081 Arabidops
3	767	59.3	1375	3 CNS0A2S8	BX827187 Arabidops
4	760.8	58.8	1312	3 CNS0A2U8	BX827470 Arabidops
5	760.4	58.8	1359	3 CNS0A2R3	BX827477 Arabidops
6	461.2	35.7	699	8 B77316	B77316 T3213TF TAM
7	399.2	30.9	691	8 BZ087060	BZ087060 lkh10d12.
8	361	27.9	709	8 BH498028	BH498028 BOHOK687F
9	355.8	27.5	1519	3 CNS0A3B5	BX826795 Arabidops
10	302	23.4	428	1 AU227210	AU227210 AU227210
11	298	23.0	424	5 BP575101	BP575101 BP575101
12	290.4	22.5	1323	3 CNS0A36P	BX826811 Arabidops
13	282.2	21.8	558	1 AV828990	AV828990 AV828990
14	273.8	21.2	446	1 AI996743	AI996743 701668184
15	252.2	19.5	782	7 CO105310	CO105310 GR_Eb003
16	249.2	19.3	822	8 BZ491815	BZ491815 BONFU42TF
17	248.4	19.2	922	8 BZ967767	BZ967767 PUDG056TD
18	245	18.9	498	1 AU226302	AU226302 AU226302
19	242	18.7	1374	3 CNS0A0BEV	BX814955 Arabidops
20	241.2	18.7	666	5 BQ157452	BQ157452 NF105A121
21	229.4	17.7	770	9 CG344371	CG344371 OG48G15TC
22	227.8	17.6	627	5 BQ155273	BQ155273 NF078F121
23	227.2	17.6	837	6 CB292743	CB292743 UCRCS01.0
24	226.2	17.5	739	7 CO105307	CO105307 GR_Eb003

25	223.6	17.3	768	1 AJ795195	AJ795195
26	222.6	17.2	444	1 AV439952	AV439952
27	218.8	16.9	905	5 BX927581	BX927581
28	217.6	16.8	630	4 BG525645	BG525645
29	214.8	16.6	1289	3 CNS0A2HI	CNS0A2HI
30	213.6	16.5	568	4 BML76948	BML76948
31	212.2	16.4	668	5 BQ704880	BQ704880
32	210.4	16.3	1299	3 CNS0A41J	CNS0A41J
33	204.6	15.8	426	5 BP666513	BP666513
34	202.8	15.7	806	4 BI310188	BI310188
35	200.8	15.5	706	1 AJ803278	AJ803278
36	200.4	15.5	780	6 CD485871	CD485871
37	200.2	15.5	701	4 BI204933	BI204933
38	200.2	15.5	809	7 CF445584	CF445584
39	199.4	15.4	591	4 BI205616	BI205616
40	199.4	15.4	742	4 BI205696	BI205696
41	199.4	15.4	742	4 BI210251	BI210251
42	199.4	15.4	839	4 BG446722	BG446722
43	198.4	15.3	726	8 BZ045790	BZ045790
44	198.2	15.3	633	1 AI777352	AI777352
45	197	15.2	589	7 CK096760	CK096760

ALIGNMENTS

RESULT 1  
CNS0A29J  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

CNS0A29J 1340 bp mRNA linear HTC 06-FEB-2004  
Arabidopsis thaliana Full-length cDNA Complete sequence from clone  
GSLTSL71ZA05 of Adult vegetative tissue of strain col-0 of  
Arabidopsis thaliana (thale cress).  
BX827536  
HTC; GSLT\_CDNA.  
Arabidopsis thaliana (thale cress)  
Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
1 (bases 1 to 1340)  
Castelli,V., Aury,J.M., Jaillon,O., Wincker,P., Clepet,C.,  
Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V.,  
Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.  
Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences:  
A Combined Approach to Evaluate and Improve Arabidopsis Genome  
Annotation  
Unpublished  
2 (bases 1 to 1340)  
Genoscope.  
Direct Submission  
Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :  
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqrefgenoscope.cns.fr  
- Web : www.genoscope.cns.fr)  
The sequences are based on single pass reads.  
Life Technologies (a division of Invitrogen) members carried out  
full-length libraries construction : Temple G.  
Genoscope members carried out sequencing and annotation : Castelli  
V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,  
Schachter V., Weissenbach J., Salanoubat M.  
URGV INRA : Clepet C., Caboche M.  
Annotation is based on the June 2003 version of the Arabidopsis  
genome released by MIPS (Munich Information center for Protein  
Sequences). 5 prime and 3 prime are assembled with Phrap.  
http://www.genoscope.cns.fr/externe/sequences/Banque\_Projet\_EF/Full  
length  
http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.  
Location/Qualifiers  
1. 1340  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/strain="Col-0"  
/db\_xref="taxon:3702"



Query Match	60.2%;	Score 778;	DB 3;	Length 1431;
Best Local Similarity	78.9%;	Ref. No. 3.6e-207;		
Matches	964;	Conservative	0;	Mismatches 230; Indels 28; Gaps 2;
Qy	15	AGTGGTCATGTGAAATAACTCAAGTAATCTATGTCAATGTGAAGCAAGATGCATCTCGTAG	74	
Db	39	ATTTCATGTGGCATATCTCAAGTACACTATTTGCATGTGAAGCAAAAGGTACCTGATGA	98	
Qy	75	TGTAGATTACTTGATTCCTTTTCGCTAACCTGT------TGTTTTGTGATGTTTTTCAG	124	
Db	99	CGTTAGATCATTTGATCCTCTTTTCGATGAATAATAAACACGAGTGTTAATTTTTTCAG	158	
Qy	125	ATCATATAACATAGAAGCAACCTAACAGGTACAGAGTACAGAGGAATGAATACCAACATGGAAA	184	
Db	159	GTGACCAAGACTTAGAAGCAAACTTTAGATCATAGGAAACT-----	202	
Qy	185	TCGAATCTTCGTCCGTACTCAATCGAAGAACTATAAGAAATGCTTCGTATTTTCCATTTT	244	
Db	203	--GAATCATTTTCAGTACCTCAACAGGAAGAACTGTAGAGGTGGCTCGGTGTCCTCATAT	260	
Qy	245	ACGTGTTCTTTGTCCTTTCGTCGCAAGCACCTTCTACAAATTTTGGGCAGAGTTTACTATG	304	
Db	261	ACGCAATCTTTGTCACTCTTCGCAACCCACTTGTACAGTTTCTGGTAGACTGTACTATG	320	
Qy	305	AAAATGSTGGGAAGAGTACATGGATGGGAACACTTGTCCAACTAATCGGCTTCCCTGTTTC	364	
Db	321	AAAATGGAGGAAAAGCACATATGTGGTAAACACTTCTTCACTCAITGGCTTCCCTGTAC	380	
Qy	365	TGTTTCTTCTCCGCTCTCTTTTCCCAACCCAAAAATCCCAACCAACAGACAGATTTTCA	424	
Db	381	TGATTTCTTCCGCTCTCTTTTCTCGAATCAGGAACCCAAATCAACAGATACAAATTTTCA	440	
Qy	425	GAAGTTCTCTTCTTCCACCAATTTCTGGATCAGTTTACATCGTTACTTGGACTATTAGTGT	484	
Db	441	GTCACTCCCTTCTTCCACCAACCTTGATCGGTTTACTTGTGCACCTGGACTGTAGTGT	500	
Qy	485	CTGCTAACTCTTATATGTCTCTCTGGTTTACTATACCTTACCAAGTTTCTACTTTCTCCC	544	
Db	501	CTGCTTATGCTTATTTGTCTGTGCACTTGGGTTGCTTTACTTTACCAGTCTCTACTTTCTCCC	560	
Qy	545	TCATCTTGGGCTCACAATTCGGCTTCTACTGCCTTCTTCTCATATTTTCTAAACCTCGAGA	604	
Db	561	TCATTTTGGGCTCACAGTTGGCTTCTACTGCCTTTTCTCATATTTCTTAACTCGCAAA	620	
Qy	605	AGTTTCAACACCTTTCAITGTGAATTTCTCTGTTTTCCTTACTATTTTCTCTGCCTCTCTCG	664	
Db	621	AGTTCACTCCTTTGATAGTCAATTTCTTGTGTTTCTTCTTACGGTTTCTCTGCCTCTCTCG	680	
Qy	665	TGTTCAACACTGATTCGGAAAAACACAGAAAGTGTCTAGAGTAAAAATATGTGATAGGGA	724	
Db	681	TGTTCAACACTGATTCGAAAAACACAACAATGTATCTAGAGTACAAATATGTGATTCGGT	740	
Qy	725	TAAATGTACCATTTGTTGCTTCTCTGGGATTTGGATTTGCTGCTATCCCTGGTACAACTGA	784	
Db	741	TCATCTGTACATTTGGTGTCTTCCGCTGGGATTTGGATTTGGTACTATCTCTGATACAACTGC	800	
Qy	785	TCCTTCAGGAAGTTTTTAAAGAAGCAAACTTTCTCAACGGTCACTGACTTGGTTCCTTTACC	844	
Db	801	TCITTCAGGAAGTTTTTACGAAAGATACATCTCTCAGCAGTCTTGGCAATTTACC	860	
Qy	845	AATCTCTAGTTGCAAGCTGTGTGTTCTCATAGGACTTTTTCGCAAGCGGGAGTGGAAAA	904	
Db	861	AGTCTCTAGTTGCTACTTGTGTGTACTCATAGGACTTTTTCGCAAGTGGAGAGTGGAGAA	920	
Qy	905	CTTTTAAACAGTGAGATGGAAAACTTCAAACTGGGGAAGTGCCTATCGTTATGACTTTGG	964	
Db	921	TTCTGCCAAGTGAGATGAGAAACTTACATATCTGGGGAAGTGTCTATATCTTGACTTTGG	980	
Qy	965	CCTCGATAGCTATTTTCTTGGCAAGTCTTACCACTTTGGCGTCTGGGAGTCTGATCTTTCAGT	1024	
Db	981	CCTCAGCAGCTATTTTCTGGCAAGTATACACTGTTGGTGTGGGATTAATCTTTCAGT	1040	
Qy	1025	CATCTTCTGTGTTCTCCAAATTCCTAATACTGCTGTGGGATTTGCCTATATAGTTCCAGTTGTAG	1084	

Qy	180	GGAAATCGAATCTTCGTCGTAACCTCAATCGAAGAACTATGAAGAAATGGCTTCGTATTTTC	239
Db	110	GGAAACTGAATCATTTTCAGTACCTCAACCAAGAACTGTAAAGAGTGGCTCCGTGCTC	169
Qy	240	CATTACGTGTTCTTGTGCTTGTGTCGCAAGCACTTCTACAAATTTTGGGACAGTTTA	299
Db	170	CATATAGCAATCTTGTGTCATCTTCTGCGAACCACTTGTCTACAGTCTGGGTAGACTGTA	229
Qy	300	CTATGAAATGGTGGGAGAGTACATGATGGGAACACTTGTCCAACTAATTCGCTTTCCC	359
Db	230	CTATGAGAAATGGGAGAAAGACATATGTGTGAACACTTCTTCAACTCATTTGCTTTCCC	289
Qy	360	TGTTCTGTTCTCTTCGCTTCTTTCGCCAACCAAAATCCCAACCAACAGAGAGA	419
Db	290	TGTACTGATTTCTTCGCTTCTTTCGCAATCAGGCAACCAAAATCAACAGATACAAA	349
Qy	420	TTTCAGAAAGTCTCTTCCTTCCACATTTCTTGGATCAGTTTACATCGTTACTGACTATT	479
Db	350	TGTGAGTCACTTCTTCCTTCCACACCTTGCATCGGTTTACTTGTGCACTGACTGCT	409
Qy	480	AGTGTCTGTAACCTTATATGCTCTCTGTTGGTTTACTATPACTTACCAGTTTCTACTTT	539
Db	410	AGTGTCTGTTATGCTTATTTGTCGAGTTGGGTTGCTTTACTTACCAGTCTCTACTTT	469
Qy	540	CTCCCTCATCTGGGCTCAAAATTTGGCTTCACTGCTTCTTCTCATATTTTCTAAACTC	599
Db	470	CTCCCTCATCTGGGCTCACAGTTGGCTTCACTGCTTCTTCTCATATTTTCTTAACTC	529
Qy	600	GCAGAAGTTCACACTTTCATTTGGAATCTCTCTTCTCTCTTACTTCTCTGCGCT	659
Db	530	GCAAAGTTTCACTCTTGTGATGATCAATCTTGTCTCTCTTACGGTTTCTCTGCGCT	589
Qy	660	CCTCGTGTCAACACTGATTCGGAACACACAGCAAAAGTGTCTAGAGTAAATATGTGAT	719
Db	590	CCTCGTGTCAACACTGATTCGGAACACACAAATGATCTAGAGTAAATATGTGAT	649
Qy	720	AGGATATATGTACCATTTGCTTCTGCTGGGATTTGGATTTGCTATCCCTGGTACA	779
Db	650	TGGTTTATCTGTATCACTTGTGCTTCTGCTGGGATTTGGATTTGCTATCTCTGATACA	709
Qy	780	ACTGATCTCAGGAGGTTTTAAAGACGAACAACTCTCAACGCTCACTGACTTGTGCGC	839
Db	710	ACTGCTCTTCAGGAAGTTTTTCAGAGACATACATCTCTCAGAGTCTCTGTACTTGTCAA	769
Qy	840	TTACCAATCTCTAGTTGCAAGCTGTGTGTTTCTCATAGGACTTTTCGCAAGCGGGAGTG	899
Db	770	TTACCAAGTCTCTAGTTGCAACTTGTGTGTTACTCATAGGACTGTTTGCAGTGGAGAGTG	829
Qy	900	GAAAACTTTAACTAGATGGAAGAACTACAACTGGGGAAAGTGCCATACGTTATGAC	959
Db	830	GAGAACTCTGCAAGTGAATGAGAACTACAACTGGGGAAAGTGTCATATATCTTTGAC	889
Qy	960	TTTGGCTCGATAGTATTTCTCGCAAGTCTACACCAATGGCGCTGCTGGGACTGATCTT	1019
Db	890	TTTGGCTCGACAGCTATTTCTGCAAGTATACACTGTTGTTGTGTGGGATTAATCTT	949
Qy	1020	TGAGTCTATCTTCTGTGTTTCTCAATTCATTAACCTGCTGTGGGATTTGCTATAGTCCAGT	1079
Db	950	CGAGTCTTCTTCTGTGTTCTCAATTCATTAACCTGCTGTGGGACTGCTATAGTCCAGT	1009
Qy	1080	TGTAGCAGTATGTTTTCATATGATATAAATGAACGCGTCAAGATCTTCTCCATCATTTT	1139
Db	1010	TGTAGCAGTATGATTTTTCATATGATATAAATGAAGTATGATCAAAAGATTTTCTCCATCATTTT	1069
Qy	1140	AGCTATCTGGGATTCATTTCTATTTGCTATCAGCACTACCTCCAGCAAAAGAGTTGAA	1199
Db	1070	AGCTATCTGGGCTTCTCTCATTTTGTCTATCAGCACTACCTCCAGCAAAAGAGTTGAA	1129
Qy	1200	GACTAGCCACAAAGTCTCTTAGGAGA	1226
Db	1130	GACTTGCCAGACAAACCTGTTGAGA	1156

RESULT 4	CNS02A2U8	1312 bp	mRNA	linear	HTC 06-FEB-2004
LOCUS	Arabidopsis thaliana Full-length cDNA Complete sequence from clone				
DEFINITION	Arabidopsis thaliana Full-length cDNA Complete sequence from clone				
	Arabidopsis thaliana (thale cress).				
ACCESSION	BX827470				
VERSION	BX827470.1	GI:42460504			
KEYWORDS	HTC; GSLT cDNA.				
SOURCE	Arabidopsis thaliana (thale cress)				
ORGANISM	Arabidopsis thaliana				
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;				
	rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.				
REFERENCE	1 (bases 1 to 1312)				
AUTHORS	Castelli, V., Aury, J.M., Jaillon, O., Wincker, P., Clepet, C.,				
	Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V.,				
	Temple, G., Caboche, M., Weissenbach, J., and Salanoubat, M.				
TITLE	Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:				
	A Combined Approach to Evaluate and Improve Arabidopsis Genome				
	Annotation				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 1312)				
AUTHORS	Genoscope.				
TITLE	Direct Submission				
JOURNAL	Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :				
	BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr				
	- Web : www.genoscope.cns.fr)				
COMMENT	The sequences are based on single pass reads.				
	Life Technologies (a division of Invitrogen) members carried out				
	full-length libraries construction : Temple G.				
	Genoscope members carried out sequencing and annotation : Castelli				
	V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,				
	Schachter V., Weissenbach J., Salanoubat M.				
	URGV INRA : Clepet C., Caboche M.				
	Annotation is based on the June 2003 version of the Arabidopsis				
	genome released by MIPS (Munich Information center for Protein				
	Sequences). 5 prime and 3 prime are assembled with Phrap.				
	http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full				
	length				
	http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.				
FEATURES	Location/Qualifiers				
source	1. .1312				
	/organism="Arabidopsis thaliana"				
	/mol_type="mRNA"				
	/strain="Col-0"				
	/db_xref="taxon:3702"				
	/clone="GSLTLS59ZF07"				
	/tissue_type="Adult vegetative tissue"				
gene	1. .1312				
	/plasmid="pCMVSPORT_6"				
	/gene="At4g18205"				
ORIGIN					
Query Match	58.8%;	Score 760.8;	DB 3;	Length 1312;	
Best Local Similarity	83.5%;	Pred. No. 2.4e-202;			
Matches	875;	Conservative	0;	Mismatches 172;	Indels 1;
Gaps	1;				
Qy	180	GGAAATCGAATCTTCGTCGTAACCTCAATCGAAGAACTATGAAGAAATGGCTTCGTATTTTC	239		
Db	124	GGAACTGAATCATTTTTCAGTACCTCAGAGAGAACTGTAAGAGTGGCTCCGTGCTC	183		
Qy	240	CATTACGTGTTCTTGTGCTTGTGTCGCAAGCACTTCTACAAATTTTGGGACAGTTTA	299		
Db	184	CATATACGCAATCTTTGTCTATCTTCTGCCAACCACTTGTCTACAGTTCGCGTAGCTGA	243		
Qy	300	CTATGAAATGGTGGGAGAGTACATGATGGGAAACACTTGTCCAACTATTCGCTTCCC	359		
Db	244	CTATGAAATGGAGGGAAGAGCATAATGTTGGTAACACTTCTTCAACTCATTTGGCTTCCC	303		
Qy	360	TGTTCTGTTCTCTTCGCTTCTTTTCCCAAAATCCCAAAATCCCAAAATCCCAAGAGAGA	419		



Db 304 TGTAAGTCTCTCCGCTCTTTCTCGAATCAGGCAACCAAAATCAACAGATACAAA 363  
Qy 420 TTTTCAAGAGTCTCTCTCCCTTCAACCAATCTTGGATCAGTTTACATCGTTACTGGACTAT 479  
Db 364 TTTTCAAGTCTCTCTCCCTTCAACCAATCTTGGATCAGTTTACATCGTTACTGGACTAT 423  
Qy 480 AGTGTCTGCTTAATCTTATGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 539  
Db 424 AGTGTCTGCTTAATCTTATGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 483  
Qy 540 CTCCCTCATCTTGGCTCAATTTGGCTTCTCACTGCTTCTCTCTCTCTCTCTCTCTCTCT 599  
Db 484 CTCCCTCATCTTGGCTCAATTTGGCTTCTCACTGCTTCTCTCTCTCTCTCTCTCTCTCT 543  
Qy 600 GCAGAGTCTCACACCTTTTCAATTTGTAATCTCTCTGCTTCTCTCTCTCTCTCTCTCTCT 659  
Db 544 GCAGAGTCTCACACCTTTTCAATTTGTAATCTCTCTGCTTCTCTCTCTCTCTCTCTCTCT 603  
Qy 660 CTTCTGCTCAACACTGATTCGGAACACAGCAAAAGTGTCTAGAGTAAATATGTGA- 718  
Db 604 CTTCTGCTCAACACTGATTCGGAACACAGCAAAAGTGTCTAGAGTAAATATGTGA- 663  
Qy 719 TAGGATATATGTAATCTTACCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 778  
Db 664 TGGGGTCTATGTAATCTTACCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 723  
Qy 779 AACTGATCTCAGGAGGTTTAAAGAGCAAAATCTTCAAGGTCCTGCTGCTGCTGCTGCTG 838  
Db 724 AACTGCTCTCAGGAGGTTTCAAGAGCAAAATCTTCAAGGTCCTGCTGCTGCTGCTGCTG 783  
Qy 839 CTTTCAATCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 898  
Db 784 ATTACCACTCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 843  
Qy 899 GGAAGTCTTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 958  
Db 844 GGAAGTCTTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 903  
Qy 959 CTTTGGCTCTGATGCTATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1018  
Db 904 CTTTGGCTCTGATGCTATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 963  
Qy 1019 TTGAGTCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1078  
Db 964 TCGAGTCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1023  
Qy 1079 TTGTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1138  
Db 1024 TTGTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1083  
Qy 1139 TAGCTATCTGGGCTTCTATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1198  
Db 1084 TAGCTATCTGGGCTTCTATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1143  
Qy 1199 AGACTAGCCACAACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1226  
Db 1144 AGACTAGCCACAACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1171

## RESULT 5

CNS0A2R3 1359 bp mRNA linear HTC 06-FEB-2004  
LOCUS Arabidopsis thaliana Full-length cDNA Complete sequence from clone  
DEFINITION GSLTSL52E07 of Adult vegetative tissue of strain col-0 of  
Arabidopsis thaliana (thale cress).  
ACCESSION BX827477.1 GI:42459927  
VERSION HTC; GSLT\_cDNA.  
KEYWORDS Arabidopsis thaliana (thale cress)  
SOURCE Arabidopsis thaliana  
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.

REFERENCE 1 (bases 1 to 1359)  
AUTHORS Castellani, V., Aury, J.M., Jallion, O., Wincker, P., Clepet, C.,  
Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V.,  
Temple, G., Caboche, M., Weissenbach, J., and Salanoubat, M.  
TITLE Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences:  
A Combined Approach to Evaluate and Improve Arabidopsis Genome  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 1359)  
AUTHORS Genoscope.  
TITLE Direct Submission  
JOURNAL Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :  
BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr  
- Web : www.genoscope.cns.fr)  
COMMENT The sequences are based on single pass reads.  
Life Technologies (a division of Invitrogen) members carried out  
full-length libraries construction : Temple G.  
Genoscope members carried out sequencing and annotation : Caetelli  
V., Aury J.M., Jallion O., Wincker P., Menard M., Cruaud C.,  
URV INRA : Clepet C., Caboche M.  
Annotation is based on the June 2003 version of the Arabidopsis  
genome released by MIPS (Munich Information center for Protein  
Sequences). 5 prime and 3 prime are assembled with Phrap.  
http://www.genoscope.cns.fr/externe/sequences/Banque\_Projet\_EF/Full  
length  
http://www.genoscope.cns.fr/cgi-bin/ggb?source=Arabidopsis.  
FEATURES  
Location/Qualifiers  
1..1359 /organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/strain="Col-0"  
/db\_xref="taxon:3702"  
/clone="GSLTSL52E07"  
/tissue\_type="Adult vegetative tissue"  
/plasmid="pCMVSPORT\_6"  
1..1359  
/gene="At4g18197"  
ORIGIN  
Query Match 58.8%; Score 760.4; DB 3; Length 1359;  
Best Local Similarity 79.9%; Pred. No. 3.1e-202;  
Matches 937; Conservative 0; Mismatches 221; Indels 14; Gaps 3;  
Qy 124 GATCATAACAACATAGAAGCAAACTTAACAGGTGAGGAGAAATCAATACCACCATGGA 183  
Db 39 GGTGACAGAACTTAGAGCAAACTTATAGATCATAGAGGTGGAATCT-----GNA 89  
Qy 184 ATCGAATCTTCGTCCGTACCTCAATCGAAGAACTATAAGAAATGGCTTCGTATTTCCATT 243  
Db 90 TCATCATCATCAGCTGTCCTCAACCCGAGAACTATAAAGGTGGCTTCGTCTCCATA 149  
Qy 244 TAGCTGTTCTTTGCTTGTCTTGGCTTGGCAACACTTTCACAAATTTGGGAGAGTTTACTAT 303  
Db 150 TAGCTGTTCTTTGCTTGTCTTGTGCGCCACTAGCTACAAATTTCTGGGTAGATTGTACTAT 209  
Qy 304 GAAATGTTGGGAGAGTACATGGATGGAACTTGTCCAACTAATCGGCTTCCTCTGTT 363  
Db 210 GAAATGAGGAAATAGCACAATATGTGTATACACTTCTTCACTCTTGGCTTCCTCTGTA 269  
Qy 364 CTGTTTCTCTTCGCTTCTTTTCCAAACCAAAATCCCAACCAAGAGCAGATTTC 423  
Db 270 CTGGTTCTCTTCGCTTCTTTTCTCGAATCAGGCAACCAAAATCAAGATACAAATTC 329  
Qy 424 AGAAGTTCTCTTCTTCCATCTTGGATCAGTTTACATCGTTTACCTGGACTATTAGTG 483  
Db 330 AGTCAGTCCCTTCTTCCACCACTTGCATCGGTTTACTTGTGCACTGGACTGCTAGTG 389  
Qy 484 TCTGTAATCTTATATGCTCTGTTGTTTACTATATCTTACCACTTCTTCTTCTTCTCC 543  
Db 390 TCCGTTATGCTTATTTGCTGAGAGGTGCTCTACTTACCAGTCTCTACTTTCTCC 449  
Qy 544 CTCATCTTGGGCTCACAAATGGCTTCTACTGCTTCTTCTCATATTTTCTTAACTCGCAG 603

```
Db 450 CTATCTGGGCTACAGTGGCCCTTCTACTGCCCTTTTCTCATATTTCCCTTAACTCGAA 509
Qy 604 AAGTTCACACCTTTCAATGTAATCTCTGTTTCTCTTACTATATTTCTCTGCCCTCCTC 663
Db 510 AAGTTCACCTCTTGTATAGTGAATCTTGTCTTCTCTCACTGTATCTCTGCTCTTCTT 569
Qy 664 GTGGTCAACACTGATTCGAAACACACAAAGTGTCTAGAGTAAATATGTGATAGGG 723
Db 570 GTGGTCAACACTGATTCAGAAACTCAACTAATGTATCTAGAGTACAGATGTGATCGGG 629
Qy 724 ATAATATGTACCATTTGCTTCTGCTGGGATTTGGATTGCTCTATCCTCTGTTCAACTG 783
Db 630 TTCAATGATCAATCGTGTCTTCTGCTGGGATTTGGAGTGTACCTATCTCTAGACAAATG 689
Qy 784 ATCTTCAGGAAGTTTTAAAGAAGCAACATTTCTCAACGGTCACTGACTTGGTCTGTAC 843
Db 690 ATCTTCAGGAAGTTTTTCAAGAAGCATACATCTCTCAGCAGTAACGGACTTGGCCATTAC 749
Qy 844 CAATCTCTAGTTCGAAGCTGTGTGTTCTCTATAGGACTTTTCGCAAGCGGGAGTGGAA 903
Db 750 CAGTCTCTAGTTCGCGGTGTGTAGTTCTCTATAGGACTTTTTCGAAGTTGAGAGTGGAA 809
Qy 904 ACTTTAAAGTGTAGATGGAATACTACAAACTGGGAAAGTGCATAGTTTATGACTTTG 963
Db 810 ACTTTGCCACGTGAGATGAGAACTACAAACTCGGAAAGTGTATATGTTTGTACTTAA 869
Qy 964 GCCTCGATAGCTATTTCTGCGCAAGTGTACACCAATTCGCGTGTGGGACTGATCTTT--G 1021
Db 870 GCCTCGGCAACTATTTTCAATGCAAGTGTACACTCTTGTGTTGTGGGATATGATCTTCG 929
Qy 1022 AGTCATCTCTGTGTTCTCCAAATTCATTAAGTGTGCTGTGGGATTTGCTATAGTTCAGTTG 1081
Db 930 AGTCATCTCTGTGTTCTCCAAATTCATTCCTCCGCTGTGGGACTGCTATGTTTCCCGTTG 989
Qy 1082 TAGCAGTCAATTTGTTTCCATGATAAATGAAGCGGTCAAGATCTTCTCCATCATTTTAG 1141
Db 990 CGGCGGTGATCGTTTCCATGATAGATGAGATGAGCATCAAAATCTTCTCCATCATTTAG 1049
Qy 1142 CTATCTGGGATTCATTTTCTATGAGAGTCTCTATCAGCACTACCTCGAGAAAGTTGAAGA 1201
Db 1050 CTATCTGGGCTTCTCTCTCAATCGTCTATCAGAACTACCTCGAGAAAGTTGAATA 1109
Qy 1202 CTAGCCACAGTCTCTGAGAGTCTCTATCAGCACTACCTCGAGAAAGTTCACAAA 1261
Db 1110 CTAGCCACAGTCTCTGAGAGTCTCTAT---CTACCTGTTGAGGAAGTTCACAAA 1166
Qy 1262 ACATACATAGTGTATGATCAAAACATATTTCC 1293
Db 1167 ACATACAAAGTGTGATCAAGCATATTTCC 1198
```

```
RESULT 6
LOCUS B77316/c 699 bp DNA linear GSS 16-JAN-1998
DEFINITION T3213TF TAMU Arabidopsis thaliana genomic clone T3213, genomic survey sequence.
ACCESSION B77316
VERSION B77316.1 GI:2773955
KEYWORDS GSS.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 699)
REFERENCE Rounsley,S.D., Field,C.E., Bass,S., Linher,K., Linher,K., and Golden,K.-Berry,K., Granger,D., Suh,E., Wible,C., Adams,M.D. and Venter,J.C.
AUTHORS A BAC End Sequence Database for Identifying Minimal Overlaps in Arabidopsis Genomic Sequencing. Update 3
JOURNAL Unpublished (1997)
COMMENT Contact: Steve Rounsley
```

```
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: rounsley@tigr.org
Seq primer: M13-21
Class: BAC ends
High quality sequence stop: 699.
FEATURES
    source
        1..699
            /organism="Arabidopsis thaliana"
            /mol_type="genomic DNA"
            /ecot_type="Columbia"
            /db_xref="taxon:3702"
            /clone="T3213"
            /sex="hermaphrodite"
            /clone_lib="TAMU"
            /notes="Vector: BelobACII; Site_1: HindIII; Site_2: HindIII; Produced by Rod Wing"
```

```
ORIGIN
Query Match 35.7%; Score 461.2; DB 8; Length 699;
Best Local Similarity 85.4%; Pred. No. 3.7e-118;
Matches 514; Conservative 0; Mismatches 88; Indels 0; Gaps 0;
Qy 625 AATTCTCTGTTTCTCTTACTATTTCTCTGCCCTCTCTGTTGTTCAACACTGATTCGAA 684
Db 698 AATTCTTTGTATCTCTTACGTTTCTTTGCCCTCTCTGTTGTTCAACACTGATTCGAA 639
Qy 685 AACACACAAAGTGTCTAGAGTAAATATGTATAGGGATAATATGATCTTGGTGTCT 744
Db 638 AACACAAATGTATCTAGAGTAAATATGTATGTTGGTTCATCTGATCATTTGGTGTCT 579
Qy 745 TCTGCTGGGATTTGATTTGCTCTATCCCTGTTCAACTGATCTCTCAGAAAGTTTAAAG 804
Db 578 TCCGCTGGGATTTGATTTGTTACTATCTCTGATACAACTGCTCTTCAGAAAGTTTCAAG 519
Qy 805 AAGCAAACTCTCAACGGTCACTGTTGGTGTCTTACCAATCTCTAGTTGCAAGTGT 864
Db 518 AAGCATACATCTTCAGCAGTCTTGGACTTGGCCAAATTTACCAGTCTCTAGTTGCAACTGT 459
Qy 865 GTGGTCTCTATAGGACTTTTCGCAAGCGGGAGTGGAAACTTTTAAACAAGTGAGATGAA 924
Db 458 GTGGTACTCATAGGACTTTTTCGAGTGGAGAGTGGAGAACTCTGCCAAGTGAGATGAGA 399
Qy 925 AACTACAAACTGGGAAAGTCCCATACGTTTATGACTTTTGGCCTCGATAGCTATTTCTGG 984
Db 398 AACTACAAACTGGGAAAGTGTATATATCTTGAATTTGGCCTCAGCAGCTATTTCTGG 339
Qy 985 CAAGTCTACACCATTTGGCGTCTGGGACTGATCTTTGAGTCACTCTCTGTTGTTCTCAAT 1044
Db 338 CAAGTATACACTGTGTGGTGTGGGATTAATCTTCAGTCTTCTCTGTGTTCTCAAT 279
Qy 1045 TCCATACTCTGTGGGATTTGCCCTATAGTTCCAGTTCTAGCAGTATTTGTTTCCATGAT 1104
Db 278 TCCATACTCTGTGGGACTGCCCTATAGTTCCAGTTCTAGCAGTATTTGTTTCCATGAT 219
Qy 1105 AAATGAACCGTCAAAAGATCTTCTCCATCATTTTACCTATCTGGGATTCATTTCAATTT 1164
Db 218 AAGATGATGATCAAAAGATTTTCTCCATCATTTTGTAGTATCTGGGGCTTCTTTCATTT 159
Qy 1165 GTCTATCAGCACTACCTCGAGAAAGTGGAGACTAGCCACTAGCCACAAAGTCTCTGTAGGA 1224
Db 158 GTCTATCAGCACTACCTCGAGAAAGTGGAGACTTGAAGACTTGGCCAGACAAACCTGTGTGAG 99
Qy 1225 GA 1226
Db 98 GA 97
```

```
RESULT 7
BZ087060
```

LOCUS	BZ087060	691 bp	DNA	linear	GSS 10-OCT-2002							
DEFINITION	lkn10d12.b1 B.oleracea002 Brassica oleracea genomic, genomic survey sequence.											
ACCESSION	BZ087060											
KEYWORDS	BZ087060.1	GI:23719662										
SOURCE	GSS.											
ORGANISM	Brassica oleracea											
REFERENCE	Brassica oleracea											
AUTHORS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.											
TITLE	1 (bases 1 to 691)											
JOURNAL	Delehaunty,K., Fewell,G., Fulton,L., McCombie,W.R., Miner,T., Nash,W., Rabinowicz,P.D. and Wilson,R.K.											
COMMENT	Whole genome shotgun reads from Brassica oleracea Unpublished (2002) Contact: Richard K. Wilson Genome Sequencing Center Washington University School of Medicine Email: submissions@wustl.edu Plate: lkh10 row: d column: 12 Seq primer: -21UPpOT forward Class: shotgun High quality sequence start: 18 High quality sequence stop: 491.											
FEATURES	Location/Qualifiers											
source	1..691											
	/organism="Brassica oleracea"											
	/mol_type="genomic DNA"											
	/db_xref="taxon:3712"											
	/clone_lib="B.oleracea002"											
	/note="Vector: pOTw13; Whole genome shotgun library from flowering buds. DNA was purified from a crude nuclear prep using Brassica oleracea T0100DH3 buds provided by Thomas Osborn at the University of Wisconsin. Genomic DNA was provided by Pablo Rabinowicz (CSHL) and the shotgun library prepared at Washington University Genome Sequencing Center."											
ORIGIN												
Query Match	30.9%; Score 399.2; DB 8; Length 691;											
Best Local Similarity	77.4%; Pred. No. 1e-100;											
Matches 497; Conservative 0; Mismatches 143; Indels 2; Gaps 1;												
Qy	492	CTCTTATATGTCCTCTCTGTGGTTTACTATATACTTACCAGTTTCTACTTTCTTCTCCCTCATCTTT	551									
Db	36	CGCTTATCTATGCAATTGGTTTACTTCTACCTTACCTGTTTCCACCTTCTCCCTCATCTTT	95									
Qy	552	GGCCTCAAAATGGCCCTTCACTGCCTTCTTCTCATATTTTCTAACTCGCAGAGTTTCAC	611									
Db	96	GGCCTCACAGCTAGCCCTTCACTGCTTCTTCTTCTTACTTCTTAACTCGCAAAGTTTCAC	155									
Qy	612	ACCTTTCATTTGTGAATTTCTGTGTTTTCCTTACTATTTTCTCTGCTCCCTCTCGTGGTCAA	671									
Db	156	TCCCTTCATAGTCAACTCTCTGTTTTCCTTACCGTCTCCTCTGCTCTCTCGTGGTTGA	215									
Qy	672	CACGTATTCGMAAACACAGCAAAAGTGCTAGAGTAAATATGTGATACGGGATAATATG	731									
Db	216	CACCGAGTCAACAGACACAATATGCTTAGACTACAGTATGTGATCGGGTTTATATG	275									
Qy	732	TACCAATGGTGCTTCTGCTGGGATGGATTGCTGCTATCCCTGGTCAAACTGATCCTCAG	791									
Db	276	CACCATCGGTGCTTTCAGCTGGGATAGACTGCTGCTATCTCTGATACAACTCCTCTTCAG	335									
Qy	792	GAAGGTTTTAAAGAACAAACATTTCTCAACGGTCACTGACTTGGTTCGCTTACCAATCTCT	851									
Db	336	GAAGGTTTTCAAGGAGCATACATCGTCAGTAGTCATGGACTTGACTATCTACCAAGTCTCT	395									
Qy	852	AGTTGCAAGCTGTGCTTCTCATAGACTTTTTCGCAAGCGGGAGTGGGAACTTTTAACTTAAC	911									
Db	396	AGTTGCAAGCTGCGGTGTTGTTCTCATTGACATTTTCGTGAGCGGAGATGGGAACTTTGCC	455									
Qy	912	AAGTGAGATGGAAACCTACAAACTGGGGAAAGTGCCATAGCTTATGACTTTTGGCCTCGAT	971									

```
Db 297 CCTCTTCAACTTCTTTTTCACAACTCAGACAAACAAACAAATCAACAGATACAAATGTCAA 356
Qy 426 AAAGTTCTCTTCTTCCATCTTGGATCAGTTTACATCGTTTACTGGACTATTAGTGTG 485
Db 357 CCAGTCCCTTCTTCTTCAACCCCTTTATTAGTTTACATGTGTCACCGACTTACTCTC 416
Qy 486 TGCTAACTCTTATATGTCCTCTGTGTGGTTTACTATCTTACCAGTTTCTACTTCTCCCT 545
Db 417 TGTCTACGCTTATCTCTATGCAATGGTTTACTCTTACCTCTGTCTTCCACCTTCTCCCT 476
Qy 546 CATCTTGGCCCTCACAAATGGCCCTTCACTGCGCTTCTTCTCATATTTCTTAAACTCGCAGAA 605
Db 477 CATCTTGGCCCTCACAGCTAGCTTCACTGCGCTTCTTCTTACTTCTTAACTCGCAAA 536
Qy 606 GTTCACACTTTCATGTGAATCTCTGTCTTCTCTTACTATTTCTCTGCGCTTCTCGT 665
Db 537 GTTCACTCCCTTCATAGTCAACTCTCTCTTCTTCTTACCGTCTCTCTGCGCTCTCGT 596
Qy 666 GGTCAACTCATTCGGGAACACAGCAAAAGTGTCTAGAGTAAATATGATAGGAT 725
Db 597 GGTGACACAGTGCACGAAGACAAACAAATGTCTCTAGACTACAGTATGTATCGGGGT 656
Qy 726 AATATGTACCATTTGTGCTTCTGTGGATGGATTTGCTGTATCCCTCGGTAC 778
Db 657 TATATGCACCATCCGTGCTTCAGCTGGATAGGACTGCTGTATCTCTGTATC 709

CNS0A3B5 1519 bp mRNA linear HTC 06-FEB-2004
Arabidopsis thaliana Full-length cDNA Complete sequence from clone
GSLTF8632D12 of Flowers and buds of strain col-0 of Arabidopsis
thaliana (thale cress) .
BX826795
HTC; GSLT cDNA.
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 1519)
Castelli, V., Aury, J.M., Jaillon, O., Wincker, P., Clepet, C.,
Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V.,
Temple, G., Caboche, M., Weissenbach, J., and Salanoubat, M.
Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences:
A Combined Approach to Evaluate and Improve Arabidopsis Genome
Annotation
Unpublished
2 (bases 1 to 1519)
Genoscope.
Direct Submission
Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqrefgenoscope.cns.fr
- Web : www.genoscope.cns.fr)
The sequences are based on single pass reads.
Life technologies (a division of invitrogen) members carried out
full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli
V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,
Schachter V., Weissenbach J., Salanoubat M.
URGV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis
genome released by MIPS (Munich information center for Protein
Sequences) . 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full
length
http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.
Location/Qualifiers
1..1519
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/strain="Col-0"

FEATURES
source
```

```
/db xref="taxon:3702"
/clone="GSLTF8632D12"
/tissue_type="Flowers and buds"
/plasmid="pCMVSPORT_6"
1..1519
/gene="At4g18210"

ORIGIN
Query Match 27.5%; Score 355.8; DB 3; Length 1519;
Best Local Similarity 61.5%; Pred. No. 2.1e-88;
Matches 691; Conservative 0; Mismatches 417; Indels 15; Gaps 7;

Qy 136 ATAGAACAACCACTTAACAGGTGAGGAGAAATGAATACCAACCACCATGGAATCGAATCTCG 195
Db 123 AAAGAACCAAAATCCACAGTTCAGATGAAGAAGAAATTCAGTTAGTACGACGCA--GCA 179
Qy 196 TCCGTACTCTCAATCGAAGAACTATAGAATAATGGCTTCGTATTTTCCAFFTA-CGTGTTCTT 254
Db 180 GAAGTATCTTACTCTAACACATACAAACGGTCTCAGGCGTGACTTTATATACATTTCTT 239
Qy 255 TGTCTTGTCTGCGAAGCACTTTCTACAATTTTGGGAGAGTTTACTATGAAAATGTGTG 314
Db 240 TGTCAATTCAGGCCAAACAGTTGCTACAATTTTGGGAGAGTATACTATGACAAACGAGG 299
Qy 315 GAAGAGTACATGG-ATCGGAACACTTGTCCAACTAAATCGGCTTCCCTGTGTTCTCTCT 373
Db 300 AACAGTAAATGGCTTAGCAACGGTAGTTCAACTTGTGGCTTCTGTTCTACTTCCAT 359
Qy 374 TCCGCTTTTCCCAACCAAAATCCCAACCAACAGAAAGCAGATTTCCAGAAAGTTCT 433
Db 360 ATTATATCTTGTTCATTTAAACACATGCAACAACTGATAGAGATGGAAGAAACCTCAC 419
Qy 434 CTTCCTTACACATCTTGGATCAGTTTACATCGTTACTGCACTATTAAGTCTGCTAACT 493
Db 420 CT-----AGGAACCGGTATGTGTCAGTAGTGTGCTTCTTGTAGGACAGATT 473
Qy 494 CTTATATGCTCTGTTGGTTTACTATACCTACAGTTTCTACTTCTCCCTCATCTTGG 553
Db 474 GCTATTTGTACTCCATTTGGACTTCTTACTTACCGTTTCTACTATTCCTGATCTGTG 533
Qy 554 CCTCACAAATGGCTTCACTGCGCTTCTTCTCATATTTTCTTAAACTCCGAGAAGTTCAAC 613
Db 534 CAICTAGTTAGCTTCAATGCTTTTTTCTCTTATTTTCTTAACTCACAAAAACTTACCC 593
Qy 614 CTTCATTTGAATCTCTGTTTCTCTTACTATTTCTCTGCGCTCTCTGCTGCTCAACA 673
Db 594 CTATCATTTTAAATTTCTTTTCTCTCTTAACTATTTTCCACCTACTTTCATTTCAATA 653
Qy 674 CTGATTCGAAAAACACAGCAAAAGTGTCTAGAGTAAATATATGATAGGGATATATGTA 733
Db 654 ATGAGGAGACAGATTCACAAAAGTTACAAAAGAGATTTTTTAAAGGTTTCATATGCA 713
Qy 734 CCATTTGCTCTTCTGCTGGGATTTGATTTGCTGCTATCCCTGGTACAACTGATCTCAGGA 793
Db 714 CCGTTGCTGCTTTTGTGTTGTTATGTTTATTTTATCCCTACAAACAGCGAGCCTTTTATA 773
Qy 794 AG-GTTTTAAAGAGCAAAACATTTCTCAACCGTCACTGCTTGGTCTTACCATCTCTA 852
Db 774 AAAGTCTCAAAGAGCAAAATTTCTTCAGAAGTTATGATATGATAATTCATCATGAATCTG 833
Qy 853 GTTCAAGCTGTGTGGTTCTCATAGGACTTTTCCAAAGCGGGAGTGGAAACCTTTAAACA 912
Db 834 ATGGCGAGTTGTGTTCCGTTGGTGGGCGCTTTTGTCTCACAATGATGGAATAATTTGAC 893
Qy 913 AGTGAGATGGAATACTACAACTGGGAAAAGTGCCATACGTTATGATCTTTGGCCCTCGATA 972
Db 894 AATGAAAGGAAAACATAAAAAAAGGAAAAGTACCTTACAATATGAACCTTAATGTGAACA 953
Qy 973 GCTATTTTCTGGCAAG--TCTACACATTTGCGCTCGTGGGAGTGCATCTTTGAGTCATCTT 1030
Db 954 GCTGTTACTTGGGCCCGGTATCTTCATCGGTGGAACAAAGAACTGATCTTCAGGCTCTCT 1013
Qy 1031 CTGTGTTCTCCAAATTCAT-AACTGCTGTGGGATTCCTATAGTTCCAGTTGTAGCAGTG 1089
```

Db 1014 CTTCTATTCTAAATGCAATAAAGCGTTTGGAACTCCCAAGTGGTCTCTATCTTGGCTGTA 1073  
 QY 1090 ATTGTTTTCATGATATAAATGAACGGCTCAAGAGATCTTCTCCATCATTTTAGCTATCTCGG 1149  
 Db 1074 ATCATTTTCCATGACAAATAAATAGGTTAAAGGTGATTTCTATGATCTCTAGCTATTGG 1133  
 QY 1150 GGATTCATTTCATTTGCTATCAGCACTACCTCGACGAAAGAGTTGAAGACTAGCCAC 1209  
 Db 1134 GGTTTAACTTCTCTATGTACCAAAAATATCTTGTATGACAAAAAATCTTGAAGAAAATCAT 1193  
 QY 1210 ACAAGTCCTGTAGGAGATCTCTATCTTACTACTCTGCTGAGGAG 1252  
 Db 1194 GAAATAACAACACAGAAATCCCTGACCCACCAAGAGAG 1236

RESULT 10  
 AU227210/c  
 LOCUS AU227210 428 bp mRNA linear EST 23-APR-2002  
 DEFINITION AU227210 RAFL14 Arabidopsis thaliana cDNA clone RAFL14-88-B12 3',  
 mRNA sequence.  
 ACCESSION AU227210  
 VERSION AU227210.1 GI:19741857  
 KEYWORDS EST  
 SOURCE Arabidopsis thaliana (thale cress)  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 REFERENCE 1 (bases 1 to 428)  
 AUTHORS Seki,M., Narusaka,M., Ishida,J., Kamiya,A., Satou,M., Nakajima,M.,  
 Akiyama,K., Enju,A., Oono,Y., Sakurai,T., Carninci,P., Kawai,J.,  
 Itoh,M., Ishii,Y., Arakawa,T., Shibata,K., Shinagawa,A.,  
 Muramatsu,M., Hayashizaki,Y. and Shinozaki,K.  
 Large scale analysis of Arabidopsis full-length cDNA  
 Unpublished (2002)  
 TITLE Contact: Motoaki Seki  
 JOURNAL Plant Functional Genomics Research Group  
 COMMENT RIKEN Genomic Sciences Center  
 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan  
 Tel: 81-298-36-4359  
 Fax: 81-298-36-9060  
 Email: msekic@rkc.riken.go.jp  
 An Arabidopsis full-length cDNA library was constructed essentially  
 as reported previously (Seki et al., 1998). cDNA cleaved with BamHI  
 and XhoI was ligated to modified Lambda PLC-1 vector (Carninci et  
 al., submitted for publication) digested with BamHI and SalI. This  
 clone is in a modified pBluescript vector. Please visit our web  
 site ([http://www.gsc.riken.go.jp/e/plant/index\\_e.html](http://www.gsc.riken.go.jp/e/plant/index_e.html)) for further  
 details.

FEATURES  
 source  
 1..428  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3702"  
 /clone="RAFL14-88-B12"  
 /tissue\_type="root"  
 /lab\_host="DH10B"  
 /clone\_lib="RAFL14"  
 /note="Site\_1: BamHI; Site\_2: SalI"

ORIGIN  
 Query Match 23.4%; Score 302; DB 1; Length 428;  
 Best Local Similarity 100.0%; Pred. No. 2e-73;  
 Matches 302; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 992 ACACATTGGCGTGGGACTGATCTTTAGTGCATCTTCTGTGTTCTCCAAATCCATAA 1051  
 Db 428 ACACATTGGCGTGGGACTGATCTTTAGTGCATCTTCTGTGTTCTCCAAATCCATAA 369  
 QY 1052 CTGCTGTGGGATGCTTATAGTTCAGTTGTAGCAGTATTTTCCATGATAAATGA 1111  
 Db 368 CTGCTGTGGGATGCTTATAGTTCAGTTGTAGCAGTATTTTCCATGATAAATGA 309

QY 1112 ACGCGTCAAAGATCTTCTCCATCATTTTAGCTATCTGGGATTCATTTTGTCTATC 1171  
 Db 308 ACGCGTCAAAGATCTTCTCCATCATTTTAGCTATCTGGGATTCATTTTGTCTATC 249  
 QY 1172 AGCACTACTCTGACCAAAAAGAGTTGAAGACTAGGCACACAAGTCTCTGTAGAGATCCTC 1231  
 Db 248 AGCACTACTCTGACCAAAAAGAGTTGAAGACTAGGCACACAAGTCTCTGTAGAGATCCTC 189  
 QY 1232 ATCTACTACTCTGAGGAAGGTACACAAACATACATAGTGTATGATCAAAACATATTT 1291  
 Db 188 ATCTACTACTCTGAGGAAGGTACACAAACATACATAGTGTATGATCAAAACATATTT 129  
 QY 1292 CC 1293  
 Db 128 CC 127

RESULT 11  
 BP575101/c  
 LOCUS BP575101 424 bp mRNA linear EST 20-JUN-2004  
 DEFINITION BP575101 RAFL14 Arabidopsis thaliana cDNA clone RAFL14-90-L09 3',  
 mRNA sequence.  
 ACCESSION BP575101  
 VERSION BP575101.1 GI:48990867  
 KEYWORDS EST  
 SOURCE Arabidopsis thaliana (thale cress)  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 REFERENCE 1 (bases 1 to 424)  
 AUTHORS Seki,M., Narusaka,M., Kamiya,A., Ishida,J., Satou,M., Sakurai,T.,  
 Nakajima,M., Enju,A., Akiyama,K., Oono,Y., Muramatsu,M.,  
 Hayashizaki,Y., Kawai,J., Carninci,P., Itoh,M., Ishii,Y.,  
 Arakawa,T., Shibata,K., Shinagawa,A. and Shinozaki,K.  
 Functional annotation of a full-length Arabidopsis cDNA collection  
 Science 296 (5565), 141-145 (2002)  
 21932900  
 11910074  
 Contact: Motoaki Seki  
 Plant Functional Genomics Research Group  
 RIKEN Genomic Sciences Center  
 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan  
 Tel: 81-298-36-4359  
 Fax: 81-298-36-9060  
 Email: msekic@rkc.riken.go.jp  
 reversed clone; Please visit our web site  
 (<http://pfweb.gsc.riken.go.jp/>) for further details.

FEATURES  
 source  
 1..424  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3702"  
 /clone="RAFL14-90-L09"  
 /tissue\_type="root"  
 /lab\_host="DH10B"  
 /clone\_lib="RAFL14"  
 /note="Site\_1: BamHI; Site\_2: SalI"

ORIGIN  
 Query Match 23.0%; Score 298; DB 5; Length 424;  
 Best Local Similarity 100.0%; Pred. No. 2.6e-72;  
 Matches 298; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 996 CATTGGCGTGGGACTGATCTTTAGTGCATCTTCTGTGTTCTCCAAATCCATACTGC 1055  
 Db 424 CATTGGCGTGGGACTGATCTTTAGTGCATCTTCTGTGTTCTCCAAATCCATACTGC 365  
 QY 1056 TGTGGATTGCTATAGTTCAGTTGTAGCAGTATTTTCCATGATAAATGAACGC 1115  
 Db 364 TGTGGATTGCTATAGTTCAGTTGTAGCAGTATTTTCCATGATAAATGAACGC 305

QY 1116 GTCAAAGATCTTCTCCATCATTTTAGCTATCTGGGATTCATTTCATTGCTCTATCAGCA 1175  
|||||  
Db  
304 GTCAAAGATCTTCTCCATCATTTTAGCTATCTGGGATTCATTTCATTGCTCTATCAGCA 245  
|||||  
QY 1176 CTACCTCGACGAAAGAAGTTGAAGACTAGCCACACAAGTCTCTGTAGGAGATCCTCATCT 1235  
|||||  
Db 244 CTACCTCGACGAAAGAAGTTGAAGACTAGCCACACAAGTCTCTGTAGGAGATCCTCATCT 185  
|||||  
QY 1236 ACTACCTCTGAGGAGGTGCACACAAATACATAGTGTATGATCAAAACATATTTC 1293  
|||||  
Db 184 ACTACCTCTGAGGAAGTGCACACAAATACATAGTGTATGATCAAAACATATTTC 127  
|||||  
RESULT 12  
CNS0A36P 1323 bp mRNA linear HTC 06-FEB-2004  
LOCUS Arabidopsis thaliana Full-length cDNA Complete sequence from clone  
DEFINITION GSJTFB642P05 of Flowers and buds of strain col-0 of Arabidopsis  
thaliana (thale cress).  
ACCESSION BX826811 GI:42460388  
VERSION BX826811.1  
KEYWORDS HTC; GSLT\_cDNA.  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsiis.  
1 (bases 1 to 1323)  
Castelli,V., Aury,J.M., Jallion,O., Wincker,P., Clepet,C.,  
Menard,M., Cruaud,C., Quekier,F., Scarpelli,C., Schachter,V.,  
Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.  
TITLE Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences:  
A Combined Approach to Evaluate and Improve Arabidopsis Genome  
Annotation  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 1323)  
AUTHORS Genoscope.  
TITLE Direct Submission  
JOURNAL Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :  
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)  
COMMENT - Web : www.genoscope.cns.fr  
The sequences are based on single pass reads.  
Life Technologies (a division of Invitrogen) members carried out  
full-length libraries construction : Temple G.  
Genoscope members carried out sequencing and annotation : Castelli  
V., Aury J.M., Jallion O., Wincker P., Menard M., Cruaud C.,  
Schachter V., Weissenbach J., Salanoubat M.  
URGV INRA : Clepet C., Caboche M.  
Annotation is based on the June 2003 version of the Arabidopsis  
genome released by MIPS (Munich Information center for Protein  
Sequences). 5 prime and 3 prime are assembled with Phrap.  
http://www.genoscope.cns.fr/externe/sequences/Banque\_Projet\_EP/Full  
length  
http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.  
FEATURES  
Location/Qualifiers  
1..1323  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/strains="Col-0"  
/db\_xref="taxon:3702"  
/clone="GSJTFB642P05"  
/tissue\_type="Flowers and buds"  
/plasmid="pCMVSPORT\_6"  
1..1323  
/gene="At4g18190"  
ORIGIN  
Query Match 22.5%; Score 290.4; DB 3; Length 1323;  
Best Local Similarity 55.7%; Pred. No. 5.1e-70;  
Matches 583; Conservative 0; Mismatches 451; Indels 12; Gaps 1;  
QY 181 GAAATCGAATCTTCGTCGGTCACTCAATCGAAGAACTATAAGAAATGGCTGTTATTTCC 240  
|||||

LOCUS AV828990 558 bp mRNA linear EST 01-APR-2002  
DEFINITION AV828990 RAFL9 Arabidopsis thaliana cDNA clone RAFL09-36-N15 5', mRNA sequence.  
ACCESSION AV828990  
VERSION AV828990.1 GI:19871050  
KEYWORDS EST  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids. 1 (bases 1 to 558)  
REFERENCE Seki, M., Narusaka, M., Ishida, J., Kamiya, A., Satou, M., Nakajima, M., Oono, Y., Sakurai, T., Carninci, P., Kawai, J., Itoh, M., Ishii, Y., Arakawa, T., Shibata, K., Shinagawa, A., Muramatsu, M., Hayashizaki, Y. and Shinozaki, K.  
AUTHORS Large scale analysis of Arabidopsis full-length cDNA (2002b)  
TITLE Unpublished (2002)  
JOURNAL Contact: Motoaki Seki  
COMMENT Plant Functional Genomics Research Group  
RIKEN Genomic Sciences Center  
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan  
Tel: 81-298-36-4359  
Fax: 81-298-36-9080  
Email: mseki@rtc.riken.go.jp  
An Arabidopsis full-length cDNA library was constructed essentially as reported previously (Seki et al., 1998). cDNA cleaved with BamHI and XhoI was ligated to modified Lambda FLC-1 vector (Carninci et al., submitted for publication) digested with BamHI and SalI. This clone is in a modified pluescript vector. Please visit our web site ([http://www.gsc.riken.go.jp/e/plant/index\\_e.html](http://www.gsc.riken.go.jp/e/plant/index_e.html)) for further details.

FEATURES  
source Location/Qualifiers  
1..558  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
/clone="RAFL09-36-N15"  
/dev\_stage="plants at various developmental stages from germination to mature seeds"  
/lab\_host="DH10B"  
/clone\_lib="RAFL9"  
/note="Site 1: BamHI; Site 2: SalI; subjected to dehydration (1, 2, 5, 10, 24 hr) and cold (1, 2, 5, 10, 24 hr) treatments"

ORIGIN  
Query Match 21.8%; Score 282.2; DB 1; Length 558;  
Best Local Similarity 79.4%; Pred. No. 7.9e-68;  
Matches 332; Conservative 0; Mismatches 86; Indels 0; Gaps 0;  
QY 180 GGAAATCGAATCTTCGTCGTAACCTCAATCGAAGACTATAGAAATGCGTTCGTAATTC 239  
DB 141 GGAATCGAATCATTTTCAGTACCTCAACGAAAGACTGTAGAGTGCTCCGTCGTC 200  
QY 240 CATTTACGTGTTCTTTGTGCTTCGTCGCAAGCACTTCTACAAATTTGGGAGAGTTTA 299  
DB 201 CATATACGCAATCTTTGTTCATCTTCGCGCAACCACTTGCTACAGTTCGGTAGACTGA 260  
QY 300 CTATGAAATGGTGGGAAGAGTACATGATGGGAACACATTTGTCACAACTAATCGGTTCC 359  
DB 261 CTATGAAATGGAGGGAAGACATATGTGGTAACACTTCTTCAACTCATTTGGGTTCC 320  
QY 360 TGTTCGTGTTCTTCCTCCGCTCTTTTCCCAACCAAAATCCCAACCAACAGAGAGCAGA 419  
DB 321 TGTACTGATCTCTTCCTCCGCTCTTTTCGAAATCAGGCAACCAAAATCAACAGATACAAA 380  
QY 420 TTTTCAGAAAGTCTCTCTTCCTTCACCAATTCGTGATCAGTTTACATCGTTACTGGACTATT 479  
DB 381 TTTTCAGTCACTCCCTTCCTTCACCAACCTTGCAATCGGTTTACTTGTGCACTGGACTGCT 440  
QY 480 AGTGTCGTAACTCTTATATATGTCCTCTGTTGGTTTACTATATCTTACCAAGTTTCTACTTT 539

DB 441 AGTGTCGTCTTAAGCTTAATTTGTCGCACTGGCTGCTTTTACTTACCANTCTCTACTTT 500  
QY 540 CTCCTCATCTTGGCTCACAATTCGCTTCACTGCTTCTTCTCTCATATATTTCTAAAC 597  
DB 501 CTNCTCATCTTGGCTCACAGTTGGCTTCACTGNCNTTTTCTCATATATTTCTTAC 558

RESULT 14  
AI996743/c  
LOCUS AI996743 446 bp mRNA linear EST 08-SEP-1999  
DEFINITION 701668184 A. thaliana, Columbia Col-0, root-1 Arabidopsis thaliana cDNA clone 701668184, mRNA sequence.  
ACCESSION AI996743  
VERSION AI996743.1 GI:5843648  
KEYWORDS EST  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids. 1 (bases 1 to 446)  
REFERENCE Chen, J., Momiyama, M., Chan, E., Mooney, M., Carroon, B., Gilliland, D., Wang, X., Hillman, J., Guegler, K., Kim, C., Doyle, M., Brzoska, P., Gorgone, G., Burns, D., Griffin, J., Mouanoutoua, M., Nguyen, D., Tan, R., Rose, M., Warren, B., Ton, B., Kastury, K., Borillo, C., Carpio, T., Policky, J., Suzuki, G., Argentine, C., Shah, S., Nobrega, A., Murry, L., Turner, C., Krikorian, S., Elder, L. and Hanson, D.  
AUTHORS Arabidopsis thaliana Gene Expression Microarray  
TITLE Unpublished (1999)  
JOURNAL Contact: David Smoller, Ph.D.  
COMMENT Genome Systems, Inc., a wholly owned subsidiary of Incyte Pharmaceuticals, Inc.  
4633 World Parkway Circle, St. Louis, MO 63134, USA  
Tel: 877-577-2733  
Fax: 314-427-3324  
Email: [service@genomesystems.com](mailto:service@genomesystems.com).

FEATURES  
source Location/Qualifiers  
1..446  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/ecotype="Col-0"  
/db\_xref="taxon:3702"  
/clone="701668184"  
/tissue\_type="root"  
/dev\_stage="4 - 7 weeks"  
/clone\_lib="A. thaliana, Columbia Col-0, root-1"  
/note="Vector: pSPORT; Site 1: NotI; Site 2: SalI; cDNA library was derived from untreated root tissue from Arabidopsis thaliana, Columbia Col-0, at 4 - 7 weeks. Plants were grown in 1:1:1 peat moss/vermiculite/perlite soil at 22 deg. C +/- 3 deg. C under constant light, and watered with fertilizer. cDNA synthesis was initiated using a NotI-oligo(dT) primer. Double-stranded cDNA was blunted, ligated to SalI adaptors, digested with NotI, size-selected, and cloned into the NotI and SalI sites of the pSPORT vector."

ORIGIN  
Query Match 21.2%; Score 273.8; DB 1; Length 446;  
Best Local Similarity 89.5%; Pred. No. 1.7e-65;  
Matches 306; Conservative 0; Mismatches 33; Indels 3; Gaps 1;  
QY 952 GTTATGACTTTGGCTCGATAGCTATTTCCTGGCAAGTCTACACATTTGGCGTCTGGGA 1011  
DB 446 GTTTTGACTTTAGCCTGGCAGCTATTTCCTGGCAAGTCTACACTCTTTGGTCTTGGGA 387  
QY 1012 CTGATCTTTGAGTCATCTTCTGTGTTCTCCCAATTCACAACTGCTGTGGGATTGCTATA 1071  
DB 386 TTGATCTTCGAGTCATCTCTGTGTTCTCCCAATTCACAACTGCTGTGGGATTGCTATA 327  
QY 1072 GTTCAGTTGAGCAGTGATTTGTTTCCATGATAAATGAACGGGTCAAAAGATCTTCTCC 1131

Db 326 GTTCAGTTGCGGCAGTCATAGTTTTTCATGATAGATGAGCGCATCAAAAATCTTTCTCC 267

Qy 1132 ATCATTTTAGCTAICTGGGATTCATTTCATTTGTCTATCAGCACTACCTCGAGAAAAG 1191

Db 266 ATTATTTTAGCTATCGCGGCTCTCTTTCAATTCGTCTATCAGCACTACCTCGAGAAAAG 207

Qy 1192 AAGTTGAAGACTAGGCACACAAAGTCTCTAGGAGATCCTCATCTACTACTACCTGCTGAGGAA 1251

Db 206 AAGTTGAATACTAGGCACACAAAGTCTCTAGGAGATCCTCATCTACTACTACCTGCTGAGGAA 150

Qy 1252 GGTCAACAACATACATAGTGTATGATCAAAAACATATTTCC 1293

Db 149 GGTCAACAACATACAAAGTGTGTGATCAAAAGCATATTTCC 108

RESULT 15

LOCUS CO105310/c

DEFINITION GR\_Eb0035G15.f GR\_Eb Gossypium raimondii cDNA clone GR\_Eb0035G15

5'\_\_mRNA sequence.

ACCESSION CO105310 GI:48803996

VERSION CO105310.1

KEYWORDS EST.

SOURCE Gossypium raimondii

ORGANISM Gossypium raimondii

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosoids; eurosids II; Malvaceae; Malvoideae; Gossypium.

1 (bases 1 to 782)

AUTHORS Kim,H., Yu,Y., Kudrna,D., Hatfield,J., Stum,D., Mueller,C., Udall,J.A., Rapp,R.A., Wendel,J.F., Rao,K., Soderlund,C. and Wing,R.A.

TITLE Global assembly of Cotton ESTs

JOURNAL Unpublished (2004)

COMMENT Contact: Rod A. Wing

Arizona Genomics Institute

The University of Arizona

Forbes Building Room 303, Tucson, AZ, 85721-0036, USA

Tel: 520 626 9595

Fax: 520 621 1259

Email: <http://genome.arizona.edu>

Plate: 0035 row: G column: 15.

FEATURES

source

1..782

/organism="Gossypium raimondii"

/mol\_type="mRNA"

/db\_xref="taxon:29730"

/clone="GR\_Eb0035G15"

/tissue\_type="floral"

/dev\_stage="3 to +3 DPA"

/lab\_host="DH10B"

/clone\_lib="GR\_Eb"

/note="Vector: pCMV.SPORR-6.1; Site 1: NotI; Site 2: EcoRV; Library made by Invitrogen with RNA supplied by Wendle lab. Directional cloned into NotI-EV. Colonies plated/picked by AGI. More glycerol clones held in -80."

ORIGIN

Query Match 19.5%; Score 252.2; DB 7; Length 782;

Best Local Similarity 63.4%; Pred. No. 2.4e-59;

Matches 386; Conservative 0; Mismatches 223; Indels 0; Gaps 0;

Qy 597 CTCGAGAGTTCACACCTTTTCATTTGTAATCTCTGTTTCTCCTTACTATTTCCTGCG 656

Db 782 CTCACAAAAGTTCACCCCTTTTCATAATAAACTCTTTGGTCTCTCCATCAATCTCTTCCAC 723

Qy 657 CCTCTCGTGGTCAACACTGATTTCGAAAAACACAGCAAAAGTGTCTAGAGTAAATAATGT 716

Db 722 CCTCTCGTATTCAAAACGATTCTCTAGGTTCCACACAGTTCAGAGCCAAATATGT 663

Qy 717 GATAGGGATAATATACCATTTGGTCTTCTGCTGGGATTTGGATTTGCTATCCCTGGT 776

Db 662 GATCGGATTTGTATGCACGGTGGCTGCTTCAGACGGGTACGGATTGATGTTGTCTCTAAC 603

Qy 777 ACAAAGTCTCCTCAGGAAGGTTTTTAAGAAGCAAAACATTTCTCAACGGTCACTGACTTGGT 836

Db 602 ACAAAGTCTCCTCAGGAAGGTTTTTAAGAAGCAAAACATTTCTCAACGGTCACTGACTTGGT 543

Qy 837 CGTTTACCAATCTCTAGTTGCAAGCTGTGTGGTTCTCATAGGACTTTTTTCGCAAGCGGGA 896

Db 542 AATCTACCAAGTCACTCGCTGCAACTTTTGATGATCAGTGTGGTCTTTTTCCGCAAGCGGGA 483

Qy 897 GTGAAAACTTTTAAAGTGAAGTGAAGAACTACAACTCGGGAAAGTGCATACGTTAT 956

Db 482 ATGGAAGAGTCTGGCGGAGAAATGGAAGGGTTTAAACTGGGTAAGTCCGCATACGTTAA 423

Qy 957 GACTTTGGCCTCGATAGCTATTTCTGGGCAAGTCTACACCATTTGGCGTCTGCGGACTGAT 1016

Db 422 CGTTTGGTTGGATTCTGTAGGTTGGCAAGTTTCTCGATTGGTGGTGGTTTGTAT 363

Qy 1017 CTTTGGTCACTCTTCTGTGTTTCTCCAAATTCATAACTGTGTGGGATTTGCTATAGTTCC 1076

Db 362 CTTGGAAGCGTCATCGCTTTTCTCAATGTGATCAGCAAGTGGGACTACCCATTGTTCC 303

Qy 1077 AGTTGTAGCAAGTGAATGTTTTCATGATAAAATGAACCGCTCAAAGATCTTCTCCATCAT 1136

Db 302 AGTGTTCGATGGTGTGTTTTCACCGATCCAAATGACTGGAATTAAGGTCAATTTCCATGTT 243

Qy 1137 TTTAGCTATCTGGGGATTTCATTTTCTATCTAGCACTACTCTCGACGAAAAAGAGTT 1196

Db 242 GTTGGCCATATGGGGAATTTGTGCTTAAGTTTACCAGCACTATCTTGTATGATCGAAAACTC 183

Qy 1197 GAAGACTAG 1205

Db 182 TGATACTGG 174

Search completed: November 1, 2004, 22:30:18

Job time : 4611.44 secs





Db	181	GCTCAAAATCTGGGTAGATTGTACTATGAAAATGGAGAAATAGACATATGTGGTAAACA	240
Qy	241	CTTCTTCAACTCATTTGGCTTCCCTGTACTGGTCTCTGTTCCGCTTCTTTTCTCGAATCAGG	300
Db	241	CTTCTTCAACTCATTTGGCTTCCCTGTACTGGTCTCTGTTCCGCTTCTTTTCTCGAATCAGG	300
Qy	301	CAACCCAAATCAACAGATACAAAATTTAGTCAGTCCCTTCCCTTACCACCCCTTGCATCG	360
Db	301	CAACCCAAATCAACAGATACAAAATTTAGTCAGTCCCTTCCCTTACCACCCCTTGCATCG	360
Qy	361	GTTTACTTGTGCACCTGGAGCTGCTAGTGCCTTATGCTTATTTGCTGCGAGTAGGGTTG	420
Db	361	GTTTACTTGTGCACCTGGAGCTGCTAGTGCCTTATGCTTATTTGCTGCGAGTAGGGTTG	420
Qy	421	CTCTACTTACCAGTCTCTTACTTCTTCCCTCATCTTGGCCTCACAGTTGGCCTTCACTGCC	480
Db	421	CTCTACTTACCAGTCTCTTACTTCTTCCCTCATCTTGGCCTCACAGTTGGCCTTCACTGCC	480
Qy	481	TTTTTCTCATATTTCTTAACTCGCAAAAGTTCTACTCTTTTGATAGTCAGTTCTTTGCTT	540
Db	481	TTTTTCTCATATTTCTTAACTCGCAAAAGTTCTACTCTTTTGATAGTCAGTTCTTTGCTT	540
Qy	541	CTCCTCACTGATCTCTGCTCTTCTTCTGCTCAACACTGATTCAGAAACTCAACTAAT	600
Db	541	CTCCTCACTGATCTCTGCTCTTCTTCTGCTCAACACTGATTCAGAAACTCAACTAAT	600
Qy	601	GTATCTAGAGTACAGTATGTGATCGGTTCTATATGTACATCGGTCTTCGGCTGGGATT	660
Db	601	GTATCTAGAGTACAGTATGTGATCGGTTCTATATGTACATCGGTCTTCGGCTGGGATT	660
Qy	661	GGACTGTTACTATCTCTGATACAAATGCTCTTTCAGGAAAGTTTTCACGAGCATACATCC	720
Db	661	GGACTGTTACTATCTCTGATACAAATGCTCTTTCAGGAAAGTTTTCACGAGCATACATCC	720
Qy	721	TCAGCAGTACCGGACTTCGCCATTTTACCAGTCTCTAGTTGGAGTTGTAGTTCTCAT	780
Db	721	TCAGCAGTACCGGACTTCGCCATTTTACCAGTCTCTAGTTGGAGTTGTAGTTCTCAT	780
Qy	781	GGACTTTTTCGAAGTGGAGTGGGAACTTTTGCCAAAGTGAAGTGAAGAACTACAACTC	840
Db	781	GGACTTTTTCGAAGTGGAGTGGGAACTTTTGCCAAAGTGAAGTGAAGAACTACAACTC	840
Qy	841	GGGAAAGTGTATATGTTTGAATTTAGCTTTAGCTTCGGCAGCTATTTCTGCGAAGTCTACCT	900
Db	841	GGGAAAGTGTATATGTTTGAATTTAGCTTTAGCTTCGGCAGCTATTTCTGCGAAGTCTACCT	900
Qy	901	CCTGCTCTTGTGGGATTGATCTTCAGTGCATCTCTGTTCTTCCAAATTCATTAACAGCT	960
Db	901	CCTGCTCTTGTGGGATTGATCTTCAGTGCATCTCTGTTCTTCCAAATTCATTAACAGCT	960
Qy	961	GTGGGATTCCTATAGTTCCAGTTTCCGCGCAGTGTAGTTTCCATGATAGAAATGGACGCA	1020
Db	961	GTGGGATTCCTATAGTTTCCAGTTTCCGCGCAGTGTAGTTTCCATGATAGAAATGGACGCA	1020
Qy	1021	TCCAAATCTTCTCCATTTATTTAGTATCTTCGGCTTCCCTTTCATTCGTTATCAGCAC	1080
Db	1021	TCCAAATCTTCTCCATTTATTTAGTATCTTCGGCTTCCCTTTCATTCGTTATCAGCAC	1080
Qy	1081	TACCTCGACGAAAGAGTGAATTAAGTACAGCAGTGTGTAGGAGATCTTTCATCTA	1140
Db	1081	TACCTCGACGAAAGAGTGAATTAAGTACAGCAGTGTGTAGGAGATCTTTCATCTA	1140
Qy	1141	CCTGTTGAGGAGGTTCACAAACATACAAAGTGTGTGATCAAGCATATTTCC	1194
Db	1141	CCTGTTGAGGAGGTTCACAAACATACAAAGTGTGTGATCAAGCATATTTCC	1194
RESULT 2			
ATT9A21			
LOCUS			
DEFINITION Arabidopsis thaliana DNA chromosome 4, BAC clone T9A21 (ESSA project).			

ACCESSION	AL021713
VERSION	AL021713.1
KEYWORDS	GI:2832689
SOURCE	Arabidopsis thaliana (thale cress)
ORGANISM	Arabidopsis thaliana
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE	1
AUTHORS	Bevan, M., Murphy, G., Ridley, P., Hudson, S., Bancroft, I., Mewes, H.W., Mayer, K.P.X., Lemcke, K. and Schueller, C.
JOURNAL	Unpublished
REFERENCE	2
AUTHORS	EU Arabidopsis sequencing, project.
TITLE	Direct Submission
JOURNAL	Submitted (22-SEP-1999) MIPS, at the Max-Planck-Institut fuer Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail: schueller@mips.biochem.mpg.de, mayer@mips.biochem.mpg.de, Project Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK, E-mail: michael.bevan@bbsrc.ac.uk
COMMENT	Information on performance of analysis and a more detailed annotation of this entry and other sequences of chromosomes 3, 4 and 5 can be viewed at: <a href="http://www.mips.biochem.mpg.de/proj/thal/">http://www.mips.biochem.mpg.de/proj/thal/</a> .
FEATURES	Location/Qualifiers
source	1. .82697
	/organism="Arabidopsis thaliana"
	/mol_type="genomic DNA"
	/variety="Columbia"
	/db_xref="taxon:3702"
	/chromosome="4"
gene	402. .1585
	/gene="T9A21.10"
CDS	join(402. .832,946. .1090,1178. .1585)
	/gene="T9A21.10"
	/notes="similarity to SPF1 protein, sweet protein, PIR2:551529 and WRKY protein family, Petroselinum crispum, MNOS:S72443, MNOS:S72444, MNOS:S72445"
	/codon_start=1
	/product="DNA binding-like protein"
	/protein_id="CAA16788.1"
	/db_xref="GI:4038481"
	/db_xref="GOA:O49720"
	/db_xref="GOA:Q8VWJ2"
	/db_xref="Swiss-Prot:Q8VWJ2"
	/translation="MSNETRLNYQYPPSSFSILHEMNLPTSPSSVGNLPQNGFNP STYFTDCIQSSPAAYELQKTFGLSPSSSEVFNSSIDQEPNRDVTNDVINGACNE TETRVGPSNSSSEADHPGDSKGRKRELVEEQIISKVGKTKKTEVKQKREPRV SFTWKSEVDHLEDGYRWKYGOKAVKNSPYPRIIANGNENRKYRCTTKCNKVRVE RSQDPTVITTYEGQHNHPIPTNLGSSAAAFSADLMTBPSFAHDMERTAAITNG GSVAALDYGQSGYGSVNSNPSHQVTHQGEYELLREIFFSIFPKQEP"
exon	402. .832
	/gene="T9A21.10"
	/number=1
intron	833. .945
	/gene="T9A21.10"
	/number=1
exon	946. .1090
	/gene="T9A21.10"
	/number=2
intron	1091. .1177
	/gene="T9A21.10"
	/number=2
exon	1178. .1595
	/gene="T9A21.10"
	/number=3
gene	4778. .6211
	/gene="T9A21.20"
CDS	join(<4778. .4921,5023. .5187,5263. .5552,5638. .5866, 5951. .6211)
	/gene="T9A21.20"
	/notes="strong similarity to exopolysgalacturonase, Arabidopsis thaliana, PIR2:S34266"



Db	10985	CTTCTTCAACTCATTTGGCTTCCCTGTACTGTTCTGTTCCGCTCTTTCTTCGAAATCAGG	11044
QY	301	CAACCCAAATCAACAGATACAAAATTTTCAGTCAGTCCCTTCCCTTACACACCCCTTGCATCG	360
Db	11045	CAACCCAAATCAACAGATACAAAATTTTCAGTCAGTCCCTTCCCTTACACACCCCTTGCATCG	11104
QY	361	GTTTACTTGTGCACCTGGACTGCTAGTGTCCGCTTATGCTTATTTGTCTGCAGTAGGGTTG	420
Db	11105	GTTTACTTGTGCACCTGGACTGCTAGTGTCCGCTTATGCTTATTTGTCTGCAGTAGGGTTG	11164
QY	421	CTCTACTTACCAGTCTCTACTTCTTCCCTCATCTTGTGGCTCACAGTTGGCTTCCACTGCC	480
Db	11165	CTCTACTTACCAGTCTCTACTTCTTCCCTCATCTTGTGGCTCACAGTTGGCTTCCACTGCC	11224
QY	481	TTTTTCTCATATTTCTTAACTCGCAAAAGTTCACTCCTTTTGATAGTCAGTCTTTTGCTT	540
Db	11225	TTTTTCTCATATTTCTTAACTCGCAAAAGTTCACTCCTTTTGATAGTCAGTCTTTTGCTT	11284
QY	541	CTCCTCACTGTATCTCTGCTCTTCTTGTGTGCTCAACACTGATTCAGAAAACTCAACTAAT	600
Db	11285	CTCCTCACTGTATCTCTGCTCTTCTTGTGTGCTCAACACTGATTCAGAAAACTCAACTAAT	11344
QY	601	GTATCTAGAGTACAGTATGTGATCGGTTTCAATATGTACCATCGTGTTCGCTGGGATT	660
Db	11345	GTATCTAGAGTACAGTATGTGATCGGTTTCAATATGTACCATCGTGTTCGCTGGGATT	11404
QY	661	GGACTGTTTACTCTCTGATACAAATGCTCTTCAGGAAAGTTTTCAGGAAGCATACATCC	720
Db	11405	GGACTGTTTACTCTCTGATACAAATGCTCTTCAGGAAAGTTTTCAGGAAGCATACATCC	11464
QY	721	TCAGCAGTCAAGGCTTGGCCATTTACAGTCTCTAGTTGGAGTTGTGTAGTTCTCATTA	780
Db	11465	TCAGCAGTCAAGGCTTGGCCATTTACAGTCTCTAGTTGGAGTTGTGTAGTTCTCATTA	11524
QY	781	GGACTTTTTCGAAGTGGAGAGTGGGAACTTTGCCAAGTGAAGAGAGAACTACAAATC	840
Db	11525	GGACTTTTTCGAAGTGGAGAGTGGGAACTTTGCCAAGTGAAGAGAGAACTACAAATC	11584
QY	841	GGGAAAGTGTCTATATTTTCACTTTAGCTTCGCGAGCTATTTCTCGCAAGTCTACACT	900
Db	11585	GGGAAAGTGTCTATATTTTCACTTTAGCTTCGCGAGCTATTTCTCGCAAGTCTACACT	11644
QY	901	CCTGGTCTTGTGGATTTGATCTTCGAGTCATCTCTGTGTTCTCCAATTCATTAACAGCT	960
Db	11645	CCTGGTCTTGTGGATTTGATCTTCGAGTCATCTCTGTGTTCTCCAATTCATTAACAGCT	11704
QY	961	GTGGGATTCCTATAGTTCCAGTTGCGGCACTGATAGTTTCCATGATAGAAATCGACGCA	1020
Db	11705	GTGGGATTCCTATAGTTCCAGTTGCGGCACTGATAGTTTCCATGATAGAAATCGACGCA	11764
QY	1021	TCCAAAATCTTCTCCATTTATTTAGCTATCTCGCGCTTCCTTTTCATTGCTATCAGCAC	1080
Db	11765	TCCAAAATCTTCTCCATTTATTTAGCTATCTCGCGCTTCCTTTTCATTGCTATCAGCAC	11824
QY	1081	TACCTCGAGGAAAGAGTGAATACTAGCCACACAGTGTCTGAGGAGATCTTCACTTA	1140
Db	11825	TACCTCGAGGAAAGAGTGAATACTAGCCACACAGTGTCTGAGGAGATCTTCACTTA	11884
QY	1141	CCTGTTGAGGAAGGTTCACAAAAACATACAAAGTGTGTGATCAAAAGCATATTTC	1194
Db	11885	CCTGTTGAGGAAGGTTCACAAAAACATACAAAGTGTGTGATCAAAAGCATATTTC	11938

RESULT 3  
ATCHRIV48  
LOCUS ATCHRIV48 194143 bp DNA linear PLN 16-MAR-2000  
DEFINITION Arabidopsis thaliana DNA chromosome 4, contig fragment No. 48.  
ACCESSION AL161548  
VERSION AL161548.2 GI:7268504  
KEYWORDS Arabidopsis thaliana (thale cress)  
SOURCE Arabidopsis thaliana  
ORGANISM Arabidopsis thaliana

Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta; Spermatophyta; Magnoliophyta, eudicotyledons; core eudicots, rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
1 (bases 1 to 111084)  
Murphy,G., Ridley,P., Hudson,S., Mewes,H.W., Lemcke,K. and Mayer,K.F.X.  
Unpublished  
2 (bases 107966 to 194143)  
Halbert,H., Braun,M., Hoizer,E., Brandt,A., Duesterhoeft,A., Mewes,H.W., Lemcke,K. and Mayer,K.F.X.  
Unpublished  
3 (bases 1 to 194143)  
EU Arabidopsis sequencing, project.  
Direct Submission  
Submitted (10-MAR-2000) MIPS, at the Max-Planck-Institut fuer Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail: lemcke@mips.biochem.mpg.de, mayer@mips.biochem.mpg.de  
Project Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK, E-mail: michael.bevan@bbsrc.ac.uk  
Information on performance of analysis and a more detailed annotation of this entry and other sequences of chromosomes 3, 4 and 5 can be viewed at: <http://www.mips.biochem.mpg.de/proj/thal/>  
this fragment has an overlap with ATCHRIV47 at the 5' end and an overlap with ATCHRIV49 at the 3' end.  
FEATURES  
Location/Qualifiers  
1..194143  
/organism="Arabidopsis thaliana"  
/mol\_type="genomic DNA"  
/variety="Columbia"  
/db\_xref="taxon:3702"  
/chromosome="4"  
3777..9293  
/gene="AT4g18120"  
join(6377..6796,7122..7238,7347..7501,7578..8456,8543..8627,8716..8784,8871..9293)  
/gene="AT4g18120"  
/note="similarity to Mei2-like protein, Arabidopsis thaliana, gb:D86122  
contains EST gb:W43784, AI998063.1"  
/codon\_start=1  
/product="putative protein"  
/protein\_id="CA878814.1"  
/db\_xref="GI:7268605"  
/db\_xref="GOA:Q9SVV9"  
/db\_xref="TrEMBL:Q9SVV9"  
/translation="MASGLDHFSGIGINMLDDGDSHPIGNMLPDDDEELFSGMLDDLNLSSLPATLDLEDYDLFGSGGLEETDPYDSLNGKFSRMGFADSNVDNTPQNIFQNGVSIAGEHPYGBHPSTLFPVNRINSVEDSELOALFEDNPSEKDVNQGLTVFNLAPSVNRDLNIFGVYGIKEQIRTPNKRHHKFEFFDVRSDAALKALNRTEIAGKRIKLHSPRGARRNNMLQMLEDDDSYSLNHVESPLASSPIGNWRNSPIDPLQSPSKSPIFGNLSPTNRIYFESMKTASVNNQDEGRFRSHLDHLFSSSYNNASHKASTPQQPSFGSVSSFGSLNSHPSHVETLSGSEFLWGSAPWPNPFPSSNRNHRFPYSAQNSLHQLHHIGAPSGFPFRSPETSMGSAFRGAGNMNAQRNLRSTSSPNFKMLSAPRSOLFTGNGSYLMPAATMVISIDDPLEDGSGNQPDNSGNQADIKTFOLDLSKIMRGEPRPTLKNTPNXYTRNMLAAIDENKSGTYDELYLPIDFNKKNVGVAFINMYSFKETIAYEFNKWDKFNSEKSLAYATKQKALIAHFQNSLSLMDRRRCQPIVFDGSESKYPVNNLLQTHDKMTIYLCVYVLMNWLTHTCUFETRSGSEFAARASNSTGFLIFLGARHTHKADSOVNHVVIASGLKRKYQLG"  
6377..6796  
/gene="AT4g18120"  
/number=1  
6797..7121  
/gene="AT4g18120"  
/number=1  
7122..7238  
/gene="AT4g18120"  
/number=2  
7239..7346  
/gene="AT4g18120"  
/number=2  
7347..7501  
/gene="AT4g18120"



|||||  
Db 39432 CAACCCAAATCAACAGATACAAATTCAGTCAGTCCCTCTTCACCAACCTTGCATCG 39491  
Qy 361 GTTTACTGTGCACTGGACTGCTAGTGCCTTATGCTTATTTGTCTGAGTAGGGTTG 420  
Db 39492 GTTTACTGTGCACTGGACTGCTAGTGCCTTATGCTTATTTGTCTGCACTAGGGTTG 39551  
Qy 421 CTCTACTTACCAGTCTCTACTTTCTCCCTCATCTTGGCCCTCACAGTTGGCTTCACATGCC 480  
Db 39552 CTCTACTTACCAGTCTCTACTTTCTCCCTCATCTTGGCCCTCACAGTTGGCTTCACATGCC 39611  
Qy 481 TTTTCTCATATTTCCCTTAACCTCGCAAAAGTTCACTCTCTTGATAGTCAGTTCTTTGCTT 540  
Db 39612 TTTTCTCATATTTCCCTTAACCTCGCAAAAGTTCACTCTCTTGATAGTCAGTTCTTTGCTT 39671  
Qy 541 CTCCTCACTGATCTCTGCTCTTCTTGTGCTCAACACTGATTCAGAAAACCTCAACTAAT 600  
Db 39672 CTCCTCACTGATCTCTGCTCTTCTTGTGCTCAACACTGATTCAGAAAACCTCAACTAAT 39731  
Qy 601 GTATCTAGAGTACAGTATGTGATCGGGTTTCATATGTACCATCGTGTCCGCTGGGATT 660  
Db 39732 GTATCTAGAGTACAGTATGTGATCGGGTTTCATATGTACCATCGTGTCCGCTGGGATT 39791  
Qy 661 GGACTGTTACTATCTCTGATACAAATGCTCTTCAGGAAAAGTTTTCAGCAAGCATACATCC 720  
Db 39792 GGACTGTTACTATCTCTGATACAAATGCTCTTCAGGAAAAGTTTTCAGCAAGCATACATCC 39851  
Qy 721 TCAGCAGTCACGACTGGGCAATTTACAGTCTCTAGTGTGGAGTTGTGATGTTCTCAT 780  
Db 39852 TCAGCAGTCACGACTGGGCAATTTACAGTCTCTAGTGTGGAGTTGTGATGTTCTCAT 39911  
Qy 781 GGACTTTTTCAGTGGAGTGGGAACTTTGCGAAGTGAGTGAAGAACTACAAACTC 840  
Db 39912 GGACTTTTTCAGTGGAGTGGGAACTTTGCGAAGTGAGTGAAGAACTACAAACTC 39971  
Qy 841 GGGAAAGTGTCAATGTTTGTGACTTTAGCCTCGGCGACTATTTCTCGCAAGTCTACACT 900  
Db 39972 GGGAAAGTGTCAATGTTTGTGACTTTAGCCTCGGCGACTATTTCTCGCAAGTCTACACT 40031  
Qy 901 CCTGGCTGTGGGATTTGATCTTCAGATCATCTCTGTGTCTCCAAATTCATTAACAGCT 960  
Db 40032 CCTGGCTGTGGGATTTGATCTTCAGATCATCTCTGTGTCTCCAAATTCATTAACAGCT 40091  
Qy 961 GTGGGATGTCCTATAGTTCAGTTTCGCGCAGTATGTTTCCATGATAGAACTGACGCA 1020  
Db 40092 GTGGGATGTCCTATAGTTCGAGTTTCGCGCAGTATGTTTCCATGATAGAACTGACGCA 40151  
Qy 1021 TCCAAAATCTTCTCCATTAATTTAGCTATCTCGGCTTCCCTTTCATTCGTCATACAGCAC 1080  
Db 40152 TCAAAAATCTTCTCCATTAATTTAGCTATCTCGGCTTCCCTTTCATTCGTCATACAGCAC 40211  
Qy 1081 TACCTCGAGAAAAGAGTTGAATTAATCTAGCCACACAAAGTGTGTAGAGATCTTTCACTTA 1140  
Db 40212 TACCTCGAGAAAAGAGTTGAATTAATCTAGCCACACAAAGTGTGTAGAGATCTTTCACTTA 40271  
Qy 1141 CCTGTTGAGGAGTCTACACAAATACAAAGTGTGTATCAAAAGCATATTTCC 1194  
Db 40272 CCTGTTGAGGAGTCTACACAAATACAAAGTGTGTATCAAAAGCATATTTCC 40325

RESULT 4  
AX507517  
LOCUS AX507517 3387 bp DNA linear PAT 27-SEP-2002  
DEFINITION Sequence 2212 from Patent WO0216655.  
ACCESSION AX507517  
VERSION AX507517.1 GI:23388754

SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
1  
REFERENCE

AUTHORS Harper, J. F., Krepes, J., Wang, X. and Zhu, T.  
TITLE Stress-regulated genes of plants, transgenic plants containing same, and methods of use  
JOURNAL Patent: WO 0216655-A 2212 28-FEB-2002;  
The Scripps Research Institute (US) ; Syngenta Participations AG (CH)  
FEATURES  
source 1. .3387  
/organism="Arabidopsis thaliana"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:3702"  
ORIGIN  
Query Match 85.1%; Score 1015.8; DB 6; Length 3387;  
Best Local Similarity 95.1%; Pred. No. 3.1e-273; Indels 54; Gaps 1;  
Matches 1081; Conservative 0; Mismatches 2;  
Qy 38 GTGACCAAGAACTTAGAAGCAAACTTATAGATCATGAGGTGGTAACTGAATCATCATCAT 97  
Db 1205 GTGACCAAGAACTTAGAAGCAAACTTATAGATCATGAGGTGGTAACTGAATCATCATCAT 1264  
Qy 98 CAGCTGTGCTCAAAACCGAGAACTATATAAAAGGTGGCTTCGTGCTCCATATACGTAATCT 157  
Db 1265 CAGCTGTGCTCAAAACCGAGAACTATATAAAAGGTGGCTTCGTGCTCCATATACGTAATCT 1324  
Qy 158 TTGTCCTCTTTTGGCCAGCACTAGCTACAAATCTGGGTAGATTTGTACTATGAAATCGAG 217  
Db 1325 TTGTCCTCTTTTGGCCAGCACTAGCTACAAATCTGGGTAGATTTGTACTATGAAATCGAG 1384  
Qy 218 GAAATAGCACATATGTGGTAACTTCTTCAACTCATTTGGCTTCCCTGTACTGTTCTGT 277  
Db 1385 GAAATAGCACATATGTGGTAACTTCTTCAACTCATTTGGCTTCCCTGTACTGTTCTGT 1444  
Qy 278 TCCGCTCTTTTCTCGAATCAGGCAACCCAAATCAACAGATACAAATTTCACTCAGTCCC 337  
Db 1445 TCCGCTCTTTTCTCGAATCAGGCAACCCAAATCAACAGATACAAATTTCACTCAGTCCC 1504  
Qy 338 CTTCCTTCCACACCTTGCATCGGTTCATTTGTGCACTGCTGCTGCTCTTCTTGTGCTCACTC 397  
Db 1505 CTTCCTTCCACACCTTGCATCGGTTCATTTGTGCACTGCTGCTGCTCTTCTTGTGCTCACTC 1564  
Qy 398 CTATTTTGTCTGCACTAGGGTGTCTTACTTACCAGTCTCTACTTTTCTCCCTCATCTTGG 457  
Db 1565 CTATTTTGTCTGCACTAGGGTGTCTTACTTACCAGTCTCTACTTTTCTCCCTCATCTTGG 1578  
Qy 458 CCTCACAGTTGGCTTCACTGCTTCTTCTCATATTTTCTTAACTCGCAAAAGTTCATCTC 517  
Db 1579 -----TTGGGCTTCACTGCTTCTTCTCATATTTTCTTAACTCGCAAAAGTTCATCTC 1630  
Qy 518 CTTTGATAGTCAGTTCTTCTTCTCTCACTGCTGCTGCTGCTGCTTCTTGTGCTCAACA 577  
Db 1631 CTTTGATAGTCAGTTCTTCTTCTCTCACTGCTGCTGCTGCTGCTTCTTGTGCTCAACA 1690  
Qy 578 CTGATTCAGAAAACTCAACTAATGTATCTAGAGTACAGTATGTGATCGGGTTCAATGTGA 637  
Db 1691 CTGATTCAGAAAACTCAACTAATGTATCTAGAGTACAGTATGTGATCGGGTTCAATGTGA 1750  
Qy 638 CCATCGGTGCTTCCGCTGGGATTTGGACTGTATCTATCTCTGATACAAATGCTCTTCAGGA 697  
Db 1751 CCATCGGTGCTTCCGCTGGGATTTGGACTGTATCTATCTCTGATACAAATGCTCTTCAGGA 1810  
Qy 698 AAGTTTTTCAGAAAGCATACATCTCAGCAGTCAGGACTTGGCCATTTACAGTCTCTAG 757  
Db 1811 AAGTTTTTCAGAAAGCATACATCTCAGCAGTCAGGACTTGGCCATTTACAGTCTCTAG 1870  
Qy 758 TTGCGAGTTGTAGTTCTCATAGGACTTTTTCAGAGTGGAGAGTGGGAACTTTTGCAA 817  
Db 1871 TTGCGAGTTGTAGTTCTCATAGGACTTTTTCAGAGTGGAGAGTGGGAACTTTTGCAA 1930  
Qy 818 GTGAGATGAGAACTACAACTCGGGAAGTGTCAATGTTTGTGACTTTAGCTTCGGCAG 877  
Db 1931 GTGAGATGAGAACTACAACTCGGGAAGTGTCAATGTTTGTGACTTTAGCTTCGGCAG 1990

```
QY 878 CTTATTTCTGGCAAGTCTACACCTCTGCTTGTGGGATTGATCTTCGAGTCACTCTCTG 937
Db 1991 CTTATTTCTGGCAAGTCTACACCTCTGCTTGTGGGATTGATCTTCGAGTCACTCTCTG 2050
QY 938 TGTCTCTCCAATTCATAACAGCTGTGGATTGCTATAGTTCAGATTGCGGAGTGATAG 997
Db 2051 TGTCTCTCCAATTCATAACAGCTGTGGATTGCTATAGTTCAGATTGCGGAGTGATAG 2110
QY 998 TTTTCCATGATGAATGGAGCATCCAAATCTTCTCCATTAATTTTATGCTATCTCGGGCT 1057
Db 2111 TTTTCCATGATGAATGGAGCATCCAAATCTTCTCCATTAATTTTATGCTATCTCGGGCT 2170
QY 1058 TCTTTTCATTCGTCTATCAGCACTACCTCGAAGAAAGATTGAATATAGCCACAA 1117
Db 2171 TCTTTTCATTCGTCTATCAGCACTACCTCGAAGAAAGATTGAATATAGCCACAA 2230
QY 1118 GTGCTGTAGGAGATCTTCACTACCTGTTGAGGAAGTCCACAAACATACAAAGTG 1174
Db 2231 GTGCTGTAGGAGATCTTCACTACCTGTTGAGGAAGTCCACAAACATACAAAGTG 2287

RESULT 5
AX652015 3387 bp DNA linear PAT 22-MAR-2003
LOCUS Sequence 908 from Patent WO0300899.
DEFINITION AX652015
ACCESSION AX652015
VERSION AX652015.1 GI:29154833
KEYWORDS
SOURCE
ORGANISM
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1
REFERENCE
AUTHORS Chang, H. S., Chen, W., Cooper, B., Glazebrook, J., Goff, S. A., Hou, Y. M.,
Katagiri, F., Quan, S., Tao, Y., Whitham, S., Xie, Z., Zhu, T. and Zou, G.
TITLE Plant genes involved in defense against pathogens
JOURNAL Patent: WO 0300898-A 908 03-JAN-2003;
Syngenta Participations AG (CH)
FEATURES
Location/Qualifiers
1..3387
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"

ORIGIN
Query Match 85.1%; Score 1015.8; DB 6; Length 3387;
Best Local Similarity 95.1%; Pred. No. 31e-273;
Matches 1081; Conservative 0; Mismatches 2; Indels 54; Gaps 1;

QY 38 GTGACCAAGCTTAGAAGCAACCTTATAGATCATGAGTGGTAACTGAATCATCATCAT 97
Db 1205 GTGACCAAGCTTAGAAGCAACCTTATAGATCATGAGTGGTAACTGAATCATCATCAT 1264
QY 98 CAGCTGTGCTCAAAACCGAGAACTATAAAGGTGGCTTCGTGTCTCCATATACGTAATCT 157
Db 1265 CAGCTGTGCTCAAAACCGAGAACTATAAAGGTGGCTTCGTGTCTCCATATACGTAATCT 1324
QY 158 TTGTCTCTTTTGGCCAGCCACTAGCTACAAATCTCGGTAGATTGACTATGAAATGGAG 217
Db 1325 TTGTCTCTTTTGGCCAGCCACTAGCTACAAATCTCGGTAGATTGACTATGAAATGGAG 1384
QY 218 GAAATAGCACATATGTGGTAACACTTCTTCAACTCATTTGGCTTCCTGTACTGTTCTGT 277
Db 1385 GAAATAGCACATATGTGGTAACACTTCTTCAACTCATTTGGCTTCCTGTACTGTTCTGT 1444
QY 278 TCGGCTTCTTTTCTGAATCAGGCAACCCAAATCAACAGATCAAAATTCAGTCAGTCCC 337
Db 1445 TCGGCTTCTTTTCTGAATCAGGCAACCCAAATCAACAGATCAAAATTCAGTCAGTCCC 1504
QY 338 CTTCTCTTACCAACCTTGCATCGGTTTACTTGTGCACTGGACTGCTAGTCCGCTTATG 397
Db 1505 CTTCTCTTACCAACCTTGCATCGGTTTACTTGTGCACTGGACTGCTAGTCCGCTTATG 1564
```

```
QY 398 CTTATTTGTCTGCAGTAGGGTTGCTCTACTTACCAGTCTCTACTTTCTCCCTCATCTTGG 457
Db 1565 CTTATTTGTCTGCA----- 1578
QY 458 CCTCACAGTTGGCCCTTCACTGCTTTTCTCATATTTCTTAATCTCGGAAAGTTGCACTC 517
Db 1579 -----TTGGCCCTTCACTGCTTTTCTCATATTTCTTAATCTCGGAAAGTTGCACTC 1630
QY 518 CTTTTCATAGTCAGTTCTTTTGGCTTCTCCACAGTATCTCTGCTCTTCTTGTGGTCAACA 577
Db 1631 CTTTTCATAGTCAGTTCTTTTGGCTTCTCCACAGTATCTCTGCTCTTCTTGTGGTCAACA 1690
QY 578 CTGATTTCAGAAACTCAACTAATATGATCTTAGAGTACAGTATGTGATCGGTTTCATATGTA 637
Db 1691 CTGATTTCAGAAACTCAACTAATATGATCTTAGAGTACAGTATGTGATCGGTTTCATATGTA 1750
QY 638 CCATCGGTGCTTCCGCTGGGATTGGACTGTGTACTATCTCTGATACAAATGCTCTTCAGGA 697
Db 1751 CCATCGGTGCTTCCGCTGGGATTGGACTGTGTACTATCTCTGATACAAATGCTCTTCAGGA 1810
QY 698 AAGTTTTTCAGGAAGCATACATCCTCAGCAGTCAACGACTTTGGCCATTTTACCAGTCTCTAG 757
Db 1811 AAGTTTTTCAGGAAGCATACATCCTCAGCAGTCAACGACTTTGGCCATTTTACCAGTCTCTAG 1870
QY 758 TTGCGAGTTGTGTAGTTCTTCATAGGACTTTTTCGAAGTGGAGTGGGAAACTTTTGCCAA 817
Db 1871 TTGCGAGTTGTGTAGTTCTTCATAGGACTTTTTCGAAGTGGAGTGGGAAACTTTTGCCAA 1930
QY 818 GTGAGATGAGAAACTACAAACTCGGGAAGTGTCTATGTGTTTGTAGCTTTAGGCTCGGAG 877
Db 1931 GTGAGATGAGAAACTACAAACTCGGGAAGTGTCTATGTGTTTGTAGCTTTAGGCTCGGAG 1990
QY 878 CTTATTTCTGCAAGTCTACACTCTGCTGCTTGTGGGATTGATCTTCAGTCACTCTCTG 937
Db 1991 CTTATTTCTGCAAGTCTACACTCTGCTGCTTGTGGGATTGATCTTCAGTCACTCTCTG 2050
QY 938 TGTCTCTCAATTCATAACAGCTGTGGGATTGCTTATAGTTCAGTTCGGCAGTGATAG 997
Db 2051 TGTCTCTCAATTCATAACAGCTGTGGGATTGCTTATAGTTCAGTTCGGCAGTGATAG 2110
QY 998 TTTTCCATGATGAATGGAGCATCCAAATCTTCTCCATTAATTTTATGCTATCTCGGGCT 1057
Db 2111 TTTTCCATGATGAATGGAGCATCCAAATCTTCTCCATTAATTTTATGCTATCTCGGGCT 2170
QY 1058 TCTTTTCATTCGTCTATCAGCACTACCTCGAAGAAAGATTGAATATAGCCACAA 1117
Db 2171 TCTTTTCATTCGTCTATCAGCACTACCTCGAAGAAAGATTGAATATAGCCACAA 2230
QY 1118 GTGCTGTAGGAGATCTTCACTACCTGTTGAGGAAGTCCACAAACATACAAAGTG 1174
Db 2231 GTGCTGTAGGAGATCTTCACTACCTGTTGAGGAAGTCCACAAACATACAAAGTG 2287
```

```
RESULT 6
AY096558
LOCUS Arabidopsis thaliana putative protein (At4g18200) mRNA, complete cds.
DEFINITION AY096558 1165 bp mRNA linear PLN 18-SEP-2002
ACCESSION AY096558.1 GI:20465496
VERSION AY096558.1
KEYWORDS FLI CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1 (bases 1 to 1165)
Yamada, K., Ban, J., Chan, M. M., Chang, C. H., Chang, B., Dale, J. M.,
Deng, J. M., Goldsmith, A. D., Lee, J. M., Onodera, C. S., Quach, H. L.,
Tang, C., Tortum, M., Wu, H. C., Yamamura, Y., Yu, G., Bowser, L.,
Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J.,
Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B.,
```





RESULT 7

AY074546

LOCUS

DEFINITION

AY074546 1428 bp mRNA linear PLN 18-SEP-2002  
Arabidopsis thaliana unknown protein (At4g18200) mRNA, complete

cds

AY074546

VERSION

KEYWORDS

SOURCE

ORGANISM

AY074546.1 GI:18491220

FLI CDNA

Arabidopsis thaliana (thale cress)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE

AUTHORS

1 (bases 1 to 1428)

Yamada, K., Liu, S.X., Sakano, H., Pham, P.K., Banh, J., Chung, M.K.,  
Goldsmith, A.D., Lee, J.M., Quach, H.L., Toriumi, M., Yu, G., Bowser, L.,  
Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J.,  
Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B.,  
Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T.,  
Satou, M., Seki, M., Shinn, P., Southwick, A., Shinozaki, K.,  
Davis, R.W., Ecker, J.R. and Theologis, A.

Arabidopsis Full Length cDNA Clones

Unpublished

2 (bases 1 to 1428)

Yamada, K., Banh, J., Chan, M.M., Chang, C.H., Chang, E., Dale, J.M.,  
Deng, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L.,  
Tang, C.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Bowser, L.,  
Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J.,  
Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B.,  
Lin, J., Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J.,  
Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A.,  
Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.

Direct Submission

TITLE

JOURNAL

COMMENT

Submitted (16-JAN-2002) Plant Gene Expression Center, 800 Buchanan  
Street, Albany, CA 94710, USA  
RIKEN Genomic Sciences Center (GSC) members carried out the  
collection and clustering of RAPL cDNAs (RAPL cDNA: 'RIKEN  
Arabidopsis Full-length cDNA'): Seki, M., Narusaka, M., Ishida, J.,  
Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,  
Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PGEC (SSP) Consortium members carried out the  
sequencing and annotation of the RAPL cDNAs: Yamada, K., Banh, J.,  
Chan, M.M., Chang, C.H., Chang, E., Dale, J.M., Deng, J.M.,  
Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C.,  
Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Bowser, L., Chen, H.,  
Cheuk, R., Jones, T., Karlin-Neumann, G., Kim, C., Lam, B., Lin, J.,  
Meyers, M.C., Miranda, M., Nguyen, M., Palm, C.J., Shinn, P.,  
Southwick, A., Davis, R.W., Ecker, J.R. and Theologis, A.

Yamada, K. (SSP/PGEC) and Seki, M. (RIKEN GSC) contributed equally to  
this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP/PGEC)  
contributed equally to this work as PIs.

Annotation is based on the January 2002 version of the Arabidopsis  
genome submitted to GenBank.

FEATURES

source

Location/Qualifiers

1..1428  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
/chromosome="4"  
/clone="RAF109-36-N15 (R18628)"  
/ecotype="Columbia"  
[NOTE="This clone is in a modified pBluescript vector  
(PLC-1) as a BamHI/XhoI insert."]

gene

5'UTR

CDS

1..1428

/gene="At4g18200"

1..71

/gene="At4g18200"

72..1205

/gene="At4g18200"  
/codon\_start=1  
/evidence=experimental  
/product="unknown protein"  
/protein\_id="AA169512.1"  
/db\_xref="GI:18491221"  
/translation="MGISQVHYCNGDQNLNLEALDHEETESFVPTKNCRWLRVSI  
YALFVFCQPLATVLCGLYVYENGKSTVVTLLQILGFPVLILFFRFSIRIROPKSDTD  
NFSQSPSTTLVASVYLCGLVAVLAVLSAVGLLYLPVSTFSLILASQLAFAPSPYF  
LNSQKFTPLIVNSLPLLVSSALIVNTDSNTNVSRYVYVIGICTIGASAGICLV  
LSLIQLFRKVPFTHKTSASVLDLANYQSLVATCVVLIGLFASGEWRTLPSENRNYKLG  
KVSYLTLASAAIPFQVTVGCVGLIFESSVSFNSITAVGLPIVPPVAVIVFHDHMD  
ASKPFIILAIWGLFVYQHYLDEKLKTCQTKVEBETQL"  
1206..1428  
/gene="At4g18200"  
misc\_difference 1412  
/gene="At4g18200"  
/note="not present in genomic sequence"

ORIGIN

Query Match 76.8%; Score 917.2; DB 8; Length 1428;  
Best Local Similarity 90.7%; Pred. No. 1.3e-245;  
Matches 992; Conservative 0; Mismatches 93; Indels 9; Gaps 1;  
QY 37 GGTGACCAAGAACTTAGAAGCAAACTTATAGATCATGAGGTGTAACCTGAATCATCATCA 96  
DB 102 GGTGACCAAGAACTTAGAAGCAAACTTCTAGATCATGAG-----GAAACTGAATCA 152  
QY 97 TCAGCTGTGCTTCAACCGAGAACTATATAAGGTGGCTTCGTCTCCATATACGTAAATC 156  
DB 153 TTTTCAGTACCTTCAACGAGAACTGTAAGAGGTGGCTCCGTGTCTCCATATACGCAATC 212  
QY 157 TTTTGTCTCTTTTGGCAGCACTACTACAAATTTCTGGGTAGATTGTACTATCAAAATGGA 216  
DB 213 TTTGTCTCTTCTGCAACCACTTCTGTAAGTTCTGGGTAGACTGTACTATGAAATGGA 272  
QY 217 GAAATAGACATATGTGTAAACACTCTTCAACTCATTTGGCTTCCCTGTACTGGTTCTG 276  
DB 273 GGGAAAGACATATGTGTAACTTCTTCAACTCATTTGGCTTCCCTGTACTGATTTCTC 332  
QY 277 TTCGGCTTCTTTTCTCGAATCAGGCAACCCAAATCAACAGATACAAATTTAGTCAGTCC 336  
DB 333 TTCGGCTTCTTTTCTCGAATCAGGCAACCCAAATCAACAGATACAAATTTAGTCAGTCC 392  
QY 337 CCTTCTTCCACCACCCTTGCATCGGTTTACTTGTGCACTGGACTGCTAGTGTCCGCTTAT 396  
DB 339 CCTTCTTCCACCACCCTTGCATCGGTTTACTTGTGCACTGGACTGCTAGTGTCTGCTTAT 452  
QY 397 GCTTATTTGTCTGCAGTAGGTTGCTCTACTTACCAGTCTCTACTTTTCTCCCTCATCTTG 456  
DB 453 GCTTATTTGTCTGCAGTAGGTTGCTTCTTACTTACCAGTCTCTACTTTTCTCCCTCATCTTG 512  
QY 457 GCCTCAGATTTGGCCCTTCACTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 516  
DB 513 GCCTCAGATTTGGCCCTTCACTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 572  
QY 517 CCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 576  
DB 573 CCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 632  
QY 577 ACTGATTCAGAAAACCTCAACTAATGATCTAGAGTACAGTATGATCGGGTTTCATATGT 636  
DB 633 ACTGATTCAGAAAACCTCAACTAATGATCTAGAGTACAGTATGATGATGATGATGATGATG 692  
QY 637 ACCATCGGTGCTTCGGCTGGGATTTGGACTGTTTACTATCTCTGATACAAATGCTCTTCAGG 696  
DB 693 ACCATCGGTGCTTCGGCTGGGATTTGGACTGTTTACTATCTCTGATACAAATGCTCTTCAGG 752  
QY 697 AAAGTTTTTCAGGAAGCATACATCTCAGCAGTCACGGACTTGGCCATTTTACAGTCTCTA 756  
DB 753 AAAGTTTTTCAGGAAGCATACATCTCAGCAGTCCTGGAGCTTGGCCAAATTACAGTCTCTA 812  
QY 757 GTTGGAGTTGTGTAGTTCTCATAGGACTTTTTCAGTGGAGAGTGGGAAATTTTGCCA 816

[illegible]

SOURCE ORGANISM	Arabidopsis thaliana (thale cress)	
REFERENCE AUTHORS	Arabidopsis thaliana Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.	
JOURNAL	1 Bevan, M., Murphy, G., Ridley, P., Hudson, S., Bancroft, I., Mewes, H.W., Mayer, K.F.X., Lemcke, K. and Schueller, C.	
REFERENCE AUTHORS	Unpublished	
JOURNAL	2 (bases 1 to 58427)	
REFERENCE TITLE	EU Arabidopsis sequencing, project.	
JOURNAL	Direct Submission	
COMMENT	Submitted (30-MAR-1999) MIPS, at the Max-Planck-Institut fuer Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail: schueller@mips.biochem.mpg.de, mayer@mips.biochem.mpg.de Project Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK, E-mail: michael.bevan@bbsrc.ac.uk	
FEATURES	Information on performance of analysis and a more detailed annotation of this entry and other sequences of chromosomes 3, 4 and 5 can be viewed at: <a href="http://www.mips.biochem.mpg.de/proj/thal/">http://www.mips.biochem.mpg.de/proj/thal/</a> .	
source	Location/Qualifiers	
	1. 58427	
	/organism="Arabidopsis thaliana"	
	/mol_type="genomic DNA"	
	/variety="Columbia"	
	/db_xref="taxon:3702"	
	/chromosome="4"	
misc_feature	1. 2548	
	/note="overlap to BAC T6K21; please refer to EMBL:AL021889 for analysis and annotation"	
gene	3596..5023	
	/gene="F15J5.10"	
gene	complement(join(3596..3646,3736..3801,3895..4020,4116..4281,4725..5023))	
	/gene="F15J5.10"	
CDS	complement(join(3596..3646,3736..3801,3895..4020,4116..4281,4725..5023))	
	/gene="F15J5.10"	
	/note="Contains Eukaryotic initiation factor 4E signature [DPYCFKHIEPKWEDPICANGKW]	
	contains EST gb:T21435, AAS86111"	
	/codon_start=1	
	/product="translation initiation factor eIF4E"	
	/protein_id="CAB53645.1"	
	/db_xref="GI:5816990"	
	/db_xref="GOA:O23252"	
	/db_xref="Swiss-Prot:O23252"	
	/translation="WAVEDTPKSVVTEAKPNSIENPIDRYHEGDDAERGEIAGGEG DGNVDESKSGVPESHLEHWTFFDNPAPVSKQTSWSSLRPVFTFTVEEFWSLY NNMKHPSLAHGADFCFKHIIIEPKWEDPICANGKGTWTFPKSKDSWLYTLALI GEQFDHDEICGAVNIRGQERISITWTKNASNEAAQVSIQKQKEFLDYNNSIGFII HEDAKLDNRNKNVTA"	
exon	complement(3596..3646)	
	/gene="F15J5.10"	
	/number=1	
intron	complement(3647..3735)	
	/gene="F15J5.10"	
exon	complement(3736..3801)	
	/gene="F15J5.10"	
	/number=2	
intron	complement(3802..3894)	
	/gene="F15J5.10"	
	/number=2	
exon	complement(3895..4020)	
	/gene="F15J5.10"	
	/number=3	
intron	complement(4021..4115)	
	/gene="F15J5.10"	
	/number=3	
exon	complement(4116..4281)	
	/gene="F15J5.10"	
	/number=4	

intron	complement(4282..4724)	
	/gene="F15J5.10"	
	/number=4	
exon	complement(4725..5023)	
	/gene="F15J5.10"	
	/number=5	
repeat_region	6767..6789	
gene	9077..113984	
	/gene="F15J5.20"	
CDS	join(9077..9301,9805..9859,9938..10113,10198..10419,10487..10725,10806..11331,11419..11692,11793..12736,12827..13093,13169..13432,13508..13984)	
	/gene="F15J5.20"	
	/note="similarity to multidrug resistance protein 1, Aspergillus fumigatus, U62933	
	Contains ABC transporters family signature, ABC transporter [MSGGQKRLAIARAI] [LSGGQKRLAIARAI], ATP/GTP-binding site motif A (P-loop), Atp_Gcp_A [GQSGSGKS] [GESGSGKS]"	
	/codon_start=1	
	/product="multidrug resistance protein/P-glycoprotein-like"	
	/protein_id="CAB53646.1"	
	/db_xref="GI:5816991"	
	/db_xref="GOA:Q9SVW6"	
	/db_xref="TrEMBL:Q9SVW6"	
	/translation="NEEKSSKNDDGNQKVSFFKLPSADKTDVLMVTGTTAAAGNGLTPQPMILIFQLINAFGTTDPDHVRVWVAVFIYLAIVTSCVAFVLQVSCWVTG ERQATIRGLYKLTIRQDIDYFDTEVTNGRMSGDTILIQAMGEKVGKFTQLL CTFLGGFAIFKGLAGVLCSCITPLIVIAGAAIMSLMKMAGQVAYABAGNVLE QTVGAIRTVVAFTEKOATEKYESKLEIAYKTVOOGLISGFLGTMALVIFCSYGLA VYGAKLIMEKGYNGGVINVIFAVLTCGMSIGOTSPSLNAPAGRAAFKWFETIKR SPKIDAVMSSGVLEDIRGDIELKDVPYRPARDVOI/PAGESLFPVPGKTVALLVQQS GSGKSTVISLERFDPSGQVLIDNIDKKLQKMRKSKGLVQSEPVLPFAITIKEN LAYKEDATDQEIIRTAELANAAPFDKLPGLDPTMVGHEGTMSGGQKRLAIARAI LNPKKILLDEATSALESERI/QDALVNLMSNRITVVVAHRLTIRTADVIAYVHQ GKIVKGTHTDMIQDPEGAYSQLVLOQSEKEATESESRPETSRLDVERSGSLSSAM RRSVRSNSSHSHSFLASNMFPGVNVNQTDEMEDENNVRHKKVSLKRLAHKNKE IPLVLGSIAMVHGTVPPIFGLLLSSINFPYEPAKLKKDHPWALIYALGLTNF VMIPVNTFFGIAGTKLIRKSMCFQKVHQRISWFDOTANSRYNFIYIINRRILY VLILIFICVLLPPLVRLRECESTASTVRSVLGDALIVQNIATVTTGLIITAFANWI LALIVLALSPFIVIQGVAQTFLTGFSAADAKAMTEEASQVANDAVSSIRTVASFCAEE KVMLDLYOQKCGPKNGVRLGLLSGAGFGFFLYCINCVCVFCVSGAGLIOIGKATFG EVKVFPAITIMAGVSTSAMAPDSNKAADSIFDLDSTPKIDSSEDEGTTLQNVNGDIEPHUSFRYMRPDVOI/PDRLCLTIPSGKTVALLVQSGSKSTVISMIERFYN PDSKILLDQVEIQTFLKLSWLRQQMLVSGEPILFNETIRSNATGKTGGATEEBIIA AAKAANAHNFISSLPQGYDTSVGBRGVQLSGGQKRLAIARAILKPKILLDEATSA LDAESRVVQDALDRVXG"	
	9077..9301	
exon	/gene="F15J5.20"	
	/number=1	
intron	9302..9804	
	/gene="F15J5.20"	
	/number=1	
exon	9805..9859	
	/gene="F15J5.20"	
	/number=2	
intron	9860..9937	
	/gene="F15J5.20"	
	/number=2	
exon	9938..10113	
	/gene="F15J5.20"	
	/number=3	
intron	10114..10197	
	/gene="F15J5.20"	
	/number=3	
exon	10198..10419	
	/gene="F15J5.20"	
	/number=4	
intron	10420..10486	
	/gene="F15J5.20"	
	/number=4	



REFERENCE 1 (bases 1 to 733)  
AUTHORS Stracke,R. and Palme,K.  
TITLE Signal Peptide Selection derived cDNAs from Arabidopsis thaliana  
JOURNAL leaves and guard cells  
Unpublished  
REFERENCE 2 (bases 1 to 733)  
AUTHORS Stracke,R. and Palme,K.  
TITLE Direct Submission  
JOURNAL Submitted (12-AUG-1998) Max-Delbrueck-Laboratorium in der  
Max-Planck-Gesellschaft, Carl-von-Linne-Weg 10, Koeln D-50829,  
Germany

FEATURES  
source Location/Qualifiers  
1..733  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
/clone="gps258"  
/note="signal peptide selection derived"  
38..>733  
/codon\_start=1  
/product="unknown"  
/protein\_id="AA060265.1"  
/db\_xref="GI:24417310"  
/translation="MGISQVHYCNGDNLEANLLDHEETSESPVQTXCKWLRVSI  
YAIWFCPLAVLGRLYIENGKSTVYVLLQLIGPVLILFRFFRIKRPKSTDT  
NFSQPSFTLASVYXCTGLLVSAYILSAVGLLYPVSTFSLILASQLAFAPESYF  
LNSQKFTPLIVNSLFLTVSSALLVNTDSENTNVSRVQYVIGFICTIGASAGIGLV  
LSLIQLLFRKVFVK"

ORIGIN  
Query Match 47.7%; Score 570; DB 8; Length 733;  
Best Local Similarity 91.3%; Pred. No. 2.5e-148;  
Matches 616; Conservative 0; Mismatches 50; Indels 9; Gaps 1;  
QY 37 GTGTACCAAGCTTGAACCAACCTTATAGATCATGTAGGTGGTAACTCAATCATCA 96  
DB 68 GTGTACCAAGCTTGAACCAACCTTATAGATCATGTAGGTGGTAACTCAATCATCA 118  
QY 97 TCAGTGTGCTCCATCAACGAGAACTATAAAGGTGCTTCGTGCTCCATATACGTAATC 156  
DB 119 TTTTCAGTACCTCAACGAGAACTGTAGAGGTGGTCCGTGCTCCATATACGTAATC 178  
QY 157 TTTTGTCTCTTTTGGCAGCCACTAGTACAAATTCCTGGGTAGATTGTACTATGAAAATGGA 216  
DB 179 TTTTGTCTCTTTTGGCAGCCACTAGTACAAATTCCTGGGTAGATTGTACTATGAAAATG 238  
QY 217 GGAATAGACATATGTGTACACTTCTTCAACTCATTTGGCTTCCCTGTACTGTTCTG 276  
DB 239 GGGAAAAGCACATATGTGGAAACACTTCTTCAACTCATTTGGCTTCCCTGTACTGTTCTC 298  
QY 277 TTTCCGCTCTTTTCTCGAATCAGGCAACCCAAATCAACAGATACAAATTTCACTCAGTCC 336  
DB 299 TTTCCGCTCTTTTCTCGAATCAGGCAACCCAAATCAACAGATACAAATTTCACTCAGTCC 358  
QY 337 CTTCTCTTCAACACCTTCATCGGTTTACTTGTGACCTGGAAGTCTAGTGTCCGCTTAT 396  
DB 359 CTTCTCTTCAACACCTTCATCGGTTTACTTGTGACCTGGAAGTCTAGTGTCTGTTAT 418  
QY 397 GCTTATTTGTCTGACAGTGGTGTCTTACTTACAGTCTCTACTTCTCCCTCATCTTG 456  
DB 419 GCTTATTTGTCTGACAGTGGTGTCTTACTTACAGTCTCTACTTCTCCCTCATCTTG 478  
QY 457 GCCTCACAGTTGGCTTCACTGCCTTTTCTCATATTTCTTAACTCGCAAAAGTTCACT 516  
DB 479 GCCTCACAGTTGGCTTCACTGCCTTTTCTCATATTTCTTAACTCGCAAAAGTTCACT 538  
QY 517 CTTTGTATAGTCAGTCTTGTGCTTCTCCTCACTGTATCTCTGCTCTTCTTGTGGTCAAC 576  
DB 539 CTTTGTATAGTCAGTCTTGTGCTTCTCCTCACTGTATCTCTGCTCTTCTGCTGCTCAAC 598  
QY 577 ACTGATTCAGAAACTCAACTAATGTATCTAGAGTACAGTATGTGATCGGGTTCATATGT 636  
DB 599 ACTGATTCAGAAACTCAACTAATGTATCTAGAGTACAGTATGTGATCGGGTTCATATGT 658

QY 637 ACCATCGGTGCTTCGCTGGGATTGGACTGTGTACTATCTCTGATACAAATGCTCTTCAGG 696  
DB 659 ACCATCGGTGCTTCGCTGGGATTGGATTGGTACTATCTCTGATACAAATGCTCTTCAGG 718  
QY 697 AAAGTTTTCACGAAG 711  
DB 719 AAAGTTTTCACGAAG 733  
RESULT 11  
BX664674/c 721 bp DNA linear STS 14-OCT-2003  
LOCUS Arabidopsis thaliana transposon insertion STS AT\_5.3478, sequence  
DEFINITION tagged site.  
ACCESSION BX664674  
VERSION BX664674.1 GI:37665166  
KEYWORDS STS; STS, sequence tagged site.  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsiis.  
REFERENCE 1  
AUTHORS Clarke,J.H., Bowles,B., Carter,J., Hart,D., McCullagh,B., Walsh,S.,  
Langham,S., LeGrys,C., Jones,J.D.G. and Bevan,M.  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 721)  
AUTHORS Clarke,J.H.  
TITLE Direct Submission  
JOURNAL Submitted (14-OCT-2003) Clarke J.H., John Innes Centre, Colney  
Lane, Norwich, NR4 7UJ, UK  
COMMENT AT denotes an activation tag dissociation transposon within a  
single line. ET an enhancer trap dissociation transposon, GT a gene  
trap dissociation transposon, MT a mis-expression enhancer trap  
dissociation transposon, SM a defective suppressor mutator  
transposon. 3 denotes a sequence derived from the 3' end of the  
transposon, 5 denotes a sequence derived from the 5' end of the  
transposon. BSRRC GARNET, ATIS project  
On-line seed stock requests: http://nasc.nott.ac.uk/ NASC stock  
code: N101051.  
FEATURES  
source Location/Qualifiers  
1..721  
/organism="Arabidopsis thaliana"  
/mol\_type="genomic DNA"  
/variant="Landsberg erecta NASC stock code NW20"  
/cultivar="Thil3xTn25"  
/db\_xref="taxon:3702"  
/clone="AL021713"  
1..721  
/standard\_name="AT\_5.3478"  
STS  
ORIGIN  
Query Match 45.7%; Score 546.2; DB 11; Length 721;  
Best Local Similarity 99.5%; Pred. No. 1.1e-141;  
Matches 548; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 644 GTGCTTCGCTGGGATTGGACTGTGTACTATCTCTGATACAAATGCTCTTCAGGAAAGTTT 703  
DB 707 GGGCTTCGCTGGGATTGGACTGTGTACTATCTCTGATACAAATGCTCTTCAGGAAAGTTT 648  
QY 704 TCACGAAGCATACATCCTCAGCAGTTCAGGACTTGGCCATTTACCAGTCTCTAGTTGGGA 763  
DB 647 TCACGAAGCATACATCCTCAGCAGTTCAGGACTTGGCCATTTACCAGTCTCTAGTTGGGA 588  
QY 764 GTTGTGTAGTTCTCATAGGACTTTTTCAGAGTGGAGAGTGGGAAACTTTGCCAAGTGAGA 823  
DB 587 GTTGTGTAGTTCTCATAGGACTTTTTCAGAGTGGAGAGTGGGAAACTTTGCCAAGTGAGA 528  
QY 824 TGAGAACTACAAACTCGGAAAGTGTATATGTTTACCTTTAGCCTTCGCGCAGCTATTT 883  
DB 527 TGAGAACTACAAACTCGGAAAGTGTATATGTTTACCTTTAGCCTTCGCGCAGCTATTT 468



RESULT 13	AX033550	1071 bp	DNA	linear	PAT 21-SEP-2000
LOCUS	AX033550	1071 bp	DNA	linear	PAT 21-SEP-2000
DEFINITION	Sequence 7 from Patent WO0049152.				
ACCESSION	AX033550				
VERSION	AX033550.1	GI:10280294			
KEYWORDS	Arabidopsis thaliana (thale cress)				
SOURCE	Arabidopsis thaliana				
ORGANISM	Arabidopsis thaliana				
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi				
AUTHORS	Andre, B., Buerkle, L., Frommer, W. B. and Gillissen, B.				
TITLE	Nucleic acids that code for a nucleobase transporter				
JOURNAL	Patent: WO 0049152-A 7 24-AUG-2000; ANDRE BRUNO (BE) ; BUERKLE LUKAS (DE) ; FROMMER WOLF B (DE) ; GILLISSEN BERND (DE)				
FEATURES	Location/Qualifiers				
source	1..1071				
	/organism="Arabidopsis thaliana"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:3702"				
ORIGIN					
	Query Match 34.0%; Score 405.8; DB 6; Length 1071;				
	Best Local Similarity 64.5%; Pred. No. 2.5e-102;				
	Matches 624; Conservative 0; Mismatches 337; Indels 6; Gaps 1;				
Qy	164 TCCTTTGCCAGCCACTAGCTACAACTTCCTCAACTCAATGGCTTCCTCTGCTACCTGGTTCTGTTCCGCT 223				
Db	47 TCATTGGCCAAATCAGTTTGCTACAACTTCGGCAGACTATATATGAAAAATGGAGGAAACA 106				
Qy	224 GCACATATGTGGTAACACTTCCTCAACTCAATGGCTTCCTCTGCTACCTGGTTCTGTTCCGCT 283				
Db	107 GCAATGGCTAGCAACGGTAGTTCAGCTTGAGGCTTTCCTATTCTACTTCCATATCATC 166				
Qy	284 TCCTTTCTCGAATCAGGCAACCCAAATCAACAGATACAAATTCAGTCAGTCCCTTCCT 343				
Db	167 TCTTGTGAGTCA-----AAACACATACAACTCAGAGAGATGGCAAAATTAACTCTCAC 220				
Qy	344 TCACCACTTCGATCGGTTTACTTGTGCACTGGACTGGCTAGTGTCCGCTTATGCTTATT 403				
Db	221 TTAGAACCGTGCAATGATTGTTACATAGTGTGGACTTCTGTGAGGACAGCTTGCTTACC 280				
Qy	404 TGCTGCGAGTAGGGTTGCTCTACTTTACCACTCTCTACTTTCTCCCTCATCTTGGCCCTCAC 463				
Db	281 TATATTCCATTGGACTGCTTTTACTACTCTGTTTCTTACCCCTTCCCTGATCTGTGCATCAC 340				
Qy	464 AGTTGGCTTCATGCGCTTTTCTCATATTTCTTAACTCGCAAAAGTTCACTCTCTTGA 523				
Db	341 AGTTAGGCTTTACCGGCTTCTCTCTTATTACTCAACTCACAACAACTTACTCTCTATCA 400				
Qy	524 TAGTCAGTCTTGTGCTTCTCCTCACTGTATCTCTGCTCTTCTTGTGTGCTCAACACTGATT 583				
Db	401 FTTTGAATTCCTTTTCTCCTCACTATATCTTCCACCCCTTCCTTGCAATTAATACGAGG 460				
Qy	584 CAGAAAACTCAACTAATGTATCTAGAGTACAGTATGTGATCGGGTTTCATATGTACCATCG 643				
Db	461 AATCAGATTCAAAAAAGTTTACAAAAGGAGAGTATGTCAAAGGTTTCGTATGCACCGTTG 520				
Qy	644 GTGCTTCGCTGGGATGGACTGTATCTCTGATACAAATGCTCTTCAGGAAAGTTT 703				
Db	521 GTGCATCTGCTGGGTTGGTCTACTCTTATCCCTCAACACAGCTAGGCTTTTCGTTAAAGTTT 580				
Qy	704 TCACGAGCATACATCTCTCAGCAGTACGCACTTGGCCATTTTACCAGTCTCTAGTTGCGA 763				
Db	581 TAAAGAGCAAACTTTCTCAGAAGTTATAAATATGATAATCTACATGAGTCTAGTGGCCA 640				
Qy	764 GTTGTGAGTTCTCATAGGACTTTTTCGAAGTGGAGTGGGAAACTTTTGCCAAAGTGAGA 823				
Db	641 GTTGTGTTAGCGTGGTGGGCTTTTGTCTAGTAGCGAGTGCAGAACTTTTGAGCAGTGAAG 700				





Qy 1157 ACACAA 1162  
|  
Db 1043 AGTCAA 1048

Search completed: November 1, 2004, 15:44:43  
Job time : 5297.13 secs

***THIS PAGE BLANK (USPTO)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 1, 2004, 02:40:28 ; Search time 678.972 Seconds  
(without alignments)  
9996.736 Million cell updates/sec

Title: US-09-913-767-4

Perfect score: 1293

Sequence: 1 ggaagtcctcttagagtgt.....tatgatcaaacatatattcc 1293

Scoring table: IDENTITY NUC

Gapop 10\_0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: N Geneseq\_23Sep04:\*
- 2: geneseqn1980s:\*
- 3: geneseqn1990s:\*
- 4: geneseqn2000s:\*
- 5: geneseqn2001as:\*
- 6: geneseqn2001bs:\*
- 7: geneseqn2002as:\*
- 8: geneseqn2002bs:\*
- 9: geneseqn2003as:\*
- 10: geneseqn2003bs:\*
- 11: geneseqn2003cs:\*
- 12: geneseqn2003ds:\*
- 12: geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1293	100.0	1293	3	Aaa97922 A. thalia
2	1152	89.1	3387	6	Abz14407 Arabidops
3	1152	89.1	3387	8	Ada68246 Arabidops
4	816.6	63.2	1194	3	Aaa97923 A. thalia
5	435.2	33.7	1071	3	Aaa97925 A. thalia
6	410	31.7	1047	6	Adg88190 A. thalia
7	410	31.7	1081	3	Aaa97924 A. thalia
8	354.4	27.4	592	10	Abx56692 Arabidops
9	294.4	22.8	1145	3	Aaa97921 A. thalia
10	294.4	22.8	1292	3	Aac51119 Arabidops
11	294.4	22.8	1295	3	Aac37289 Arabidops
12	289	22.4	2660	3	Aac47974 Arabidops
13	246.6	19.1	8398	6	Abn85767 Arabidops
14	230.8	17.8	1403	3	Aac37652 Arabidops
15	227.2	17.6	2175	8	Ada70748 Rice gene
16	214.4	16.6	1152	8	Ada70213 Rice gene
17	211.6	16.4	492	3	Aac36845 Arabidops
18	157	12.1	1049	3	Aaa97920 A. thalia
19	140.6	10.9	1155	8	Ada69686 Rice gene
20	128	9.9	386	6	Abq85215 Arabidops
21	126.8	9.8	1225	3	Aaa97919 A. thalia

22	126	9.7	2586	3	AAC44184	Aac44184 Arabidops
23	116	9.0	1040	8	ADA69568	Ada69568 Rice gene
24	115.2	8.9	332	12	ADP92185	Adp92185 Cotton ex
25	89.4	6.9	5520	5	AAC68924	Aac68924 FLR1 rece
26	69.2	5.4	1173	12	ADN72520	Adn72520 Thale cre
27	67.8	5.2	774	6	ABN99127	Abn99127 Arabidops
28	67	5.2	477	3	AAC37273	Aac37273 Arabidops
29	62.6	4.8	1200	8	ADA70627	Ada70627 Rice gene
30	51.8	4.0	493	3	AAC36831	Aac36831 Arabidops
31	49.2	3.8	1418	3	AAC46395	Aac46395 Arabidops
32	47.6	3.7	1421	3	AAC39779	Aac39779 Arabidops
33	46.6	3.6	446	3	AAC37188	Aac37188 Arabidops
34	46	3.6	2613	10	ADG32560	Adg32560 Generic m
35	45.6	3.5	174424	6	ABL68122	Ab168122 Ovary can
36	45.6	3.5	181343	12	ADQ19573	Adq19573 Human sof
37	45.2	3.5	2000	8	ADA71938	Ada71938 Rice gene
38	43.6	3.4	2613	10	ADG32563	Adg32563 Generic h
39	41.6	3.2	17934	6	ABL33718	Ab133718 Human inm
40	41.4	3.2	300	2	AAZ14317	Aaz14317 Human gen
41	41.2	3.2	97835	6	ABK84796	Abk84796 Human CDR
42	41	3.2	31670	8	ADA56093	Ada56093 Human CCR
43	41	3.2	31670	9	ADA02455	Ada02455 Human CCR
44	41	3.2	31670	10	ADB72194	Adb72194 Human CCR
45	40.6	3.1	543	12	ACH70417	Ach70417 Human gen

ALIGNMENTS

RESULT 1  
AAA97922  
ID AAA97922 standard; DNA; 1293 BP.  
AC AAA97922;  
XX  
DT 19-JAN-2001 (first entry)  
XX  
DE A. thaliana PUP1 DNA #4.  
XX  
KW PUP1; transgenic plant; nucleobase transporter; apical dominance;  
KW flowering behaviour; senescence; pesticide distribution; ds.  
XX  
OS Arabidopsis thaliana.  
XX  
PN DE19907209-A1.  
XX  
PD 24-AUG-2000.  
XX  
PF 19-FEB-1999; 99DE-01007209.  
XX  
PR 19-FEB-1999; 99DE-01007209.  
XX  
(FROM/) FROMMER W.  
XX  
Gillissen B, Buerkle L, Andre B, Frommer WB;  
PI WPI; 2000-566202/53.

Nucleic acid, useful for producing transgenic plants with altered nucleobase transporter, encodes a nucleobase transporter protein of Arabidopsis thaliana.  
Claim 1f; Page 13; 24pp; German.  
This invention describes a novel nucleic acid encoding a plant nucleobase transporter (I). (I) is produced by complementation of a nucleobase transport (NBT)-defective host cell with a plant gene bank by selection of NBT-positive cells. (I) is used to isolate homologous sequences from bacteria, fungi, plants, animals and humans, for expression of the encoded protein (II) in prokaryotic or eukaryotic cells, for inhibiting expression of (II) (when in antisense orientation), and to produce transgenic crop plants. The transgenic plants have modified nucleobase transport properties, e.g. altered affinity and substrate specificity

CC that may result in more efficient nucleobase transport in leaves, changes  
CC in apical dominance, flowering behaviour and senescence, or improved  
CC distribution of pesticides. This sequence encodes the Arabidopsis  
CC thaliana PUP1 protein which is described in the method of the invention  
XX  
SQ Sequence 1293 BP; 345 A; 288 C; 247 G; 413 T; 0 U; 0 Other;

Query Match 100.0%; Score 1293; DB 3; Length 1293;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1293; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	GGAAGTCTCTTAGAGTGGTCAATGGAATACTCAAGTAATCTATGTCAATGGTAAGCAA	60
Db	1	GGAAGTCTCTTAGAGTGGTCAATGGAATACTCAAGTAATCTATGTCAATGGTAAGCAA	60
Qy	61	GATGATCTCGTAGTGTAGATTACTTGATCTTTTCGCTAACCTGTGTTTTCGATGTTT	120
Db	61	GATGATCTCGTAGTGTAGATTACTTGATCTTTTCGCTAACCTGTGTTTTCGATGTTT	120
Qy	121	TCAGATCATACACATAGAGCAAACTTAACAGCTCAGGAGGAATGAATACCACCATG	180
Db	121	TCAGATCATACACATAGAGCAAACTTAACAGCTCAGGAGGAATGAATACCACCATG	180
Qy	181	GAATCGAATCTTCGTCCGTACCTCAATCGAAGAACTATAGAAATGGCTTCGTATTTC	240
Db	181	GAATCGAATCTTCGTCCGTACCTCAATCGAAGAACTATAGAAATGGCTTCGTATTTC	240
Qy	241	ATTTACGTTGTTTGTGCTTGTGCGCAAGCACTTTCTACAATTTTGGGCGAGTTTAC	300
Db	241	ATTTACGTTGTTTGTGCTTGTGCGCAAGCACTTTCTACAATTTTGGGCGAGTTTAC	300
Qy	301	TATGAATGTTGGGAGAGTACATGGATGGAGACCTTGTCCAACTAACTCGGCTCCCT	360
Db	301	TATGAATGTTGGGAGAGTACATGGATGGAGACCTTGTCCAACTAACTCGGCTCCCT	360
Qy	361	GTTCGTTTCTTCCTCGCTCTTTTCCAAACCAAAATCCAAACCAAGAGCAGAT	420
Db	361	GTTCGTTTCTTCCTCGCTCTTTTCCAAACCAAAATCCAAACCAAGAGCAGAT	420
Qy	421	TTCAAGAAAGTCTCTTCCTCCACCATCTTGGATCAGTTTACATCGTTACTGGACTATTA	480
Db	421	TTCAAGAAAGTCTCTTCCTCCACCATCTTGGATCAGTTTACATCGTTACTGGACTATTA	480
Qy	481	GTGCTGCTACTCTTATATGCTCTCTGTTGGTTTACTATCTTACCAGTTTCTACTTTC	540
Db	481	GTGCTGCTACTCTTATATGCTCTCTGTTGGTTTACTATCTTACCAGTTTCTACTTTC	540
Qy	541	TCCCTCATCTTGGCTCACAATTTGGCTTCACTGCTCTTCTCATATTTTCTAAACTCG	600
Db	541	TCCCTCATCTTGGCTCACAATTTGGCTTCACTGCTCTTCTCATATTTTCTAAACTCG	600
Qy	601	CAGAAGTTTACACCTTTTCATTTGTAATCTCTGTTTCTCTTACTATTTCTGCGCTC	660
Db	601	CAGAAGTTTACACCTTTTCATTTGTAATCTCTGTTTCTCTTACTATTTCTGCGCTC	660
Qy	661	CTCGTGTCAACACTGATTCGGAACACACGAAAGTGTCTAGATTAATATGTGATA	720
Db	661	CTCGTGTCAACACTGATTCGGAACACACGAAAGTGTCTAGATTAATATGTGATA	720
Qy	721	GGGATAATGTACCATTTGGTCTCTGCTGGGATTTGGATTTGCTATCTCGGTACAA	780
Db	721	GGGATAATGTACCATTTGGTCTCTGCTGGGATTTGGATTTGCTATCTCGGTACAA	780
Qy	781	CTGATCTCTCAGAGGTTTTAAAGCAAAATCTCTCAACGGTCACTGACTTGGTGGCT	840
Db	781	CTGATCTCTCAGAGGTTTTAAAGCAAAATCTCTCAACGGTCACTGACTTGGTGGCT	840
Qy	841	TACCAATCTTAGTGGCAAGTGTGTTTCTCATAGGACTTTTCGCAAGCGGGAGTGG	900
Db	841	TACCAATCTTAGTGGCAAGTGTGTTTCTCATAGGACTTTTCGCAAGCGGGAGTGG	900
Qy	901	AAACCTTTAAACAGTGAGATGGAAACCTPACAAACTGGGGAAGTGGCCATACGTTATGACT	960
Db	901	AAACCTTTAAACAGTGAGATGGAAACCTPACAAACTGGGGAAGTGGCCATACGTTATGACT	960

Db	901	AAACCTTTAAACAGTGAGATGGAAACCTACAAACTGGGGAAGTGGCCATACGTTATGACT	960
Qy	961	TTGGCTCTCGATAGCTATTTCTCGCAAGTCTACACCAATTGGCGTCTGGGACTGATCTTT	1020
Db	961	TTGGCTCTCGATAGCTATTTCTCGCAAGTCTACACCAATTGGCGTCTGGGACTGATCTTT	1020
Qy	1021	GAGTCATCTTCTGTGTTCTCCAATTTCCATAAATGCTGTGGATTGCTATAGTTCAGTT	1080
Db	1021	GAGTCATCTTCTGTGTTCTCCAATTTCCATAAATGCTGTGGATTGCTATAGTTCAGTT	1080
Qy	1081	GTAGCAGTGAATGTTTTCATGATGATAAAATGAACCGCTCAAAAGATCTTCTCCATCATTTA	1140
Db	1081	GTAGCAGTGAATGTTTTCATGATGATAAAATGAACCGCTCAAAAGATCTTCTCCATCATTTA	1140
Qy	1141	GCTATCTGGGGAATTCATTTTCTATCAGCACTACCTCGACGAAAGAGTTGAAG	1200
Db	1141	GCTATCTGGGGAATTCATTTTCTATCAGCACTACCTCGACGAAAGAGTTGAAG	1200
Qy	1201	ACTAGCCACACAAGTCTGTAGGAGATCCTCATCTACTGCTGAGGAAGGTTCACACA	1260
Db	1201	ACTAGCCACACAAGTCTGTAGGAGATCCTCATCTACTGCTGAGGAAGGTTCACACA	1260
Qy	1261	AACATACATAGTGTATGATCAAAACATATTTTC	1293
Db	1261	AACATACATAGTGTATGATCAAAACATATTTTC	1293

RESULT 2  
ABZ14407  
ID ABZ14407 standard; DNA; 3387 BP.

AC ABZ14407;  
XX  
DT 21-JAN-2003 (first entry)  
XX  
DE Arabidopsis thaliana stress regulated gene SEQ ID NO 2212.  
XX  
KW Arabidopsis thaliana; plant; gene; stress; transgenic; ds.  
XX  
OS Arabidopsis thaliana.  
XX  
PN WO200216655-A2.  
XX  
PD 28-FEB-2002.  
XX  
PF 24-AUG-2001; 2001WO-US026685.  
XX  
PR 24-AUG-2000; 2000US-0227866P.  
PR 26-JAN-2001; 2001US-0264647P.  
PR 22-JUN-2001; 2001US-0300111P.  
XX

(SCRI ) SCRIPPS RES INST.  
(SYGN ) SYNGENTA PARTICIPATIONS AG.

Harper JF, Krepis J, Wang X, Zhu T;

WPI; 2002-304127/34.

Identifying a stress condition to which a plant cell has been exposed and producing plants with increased tolerance to these abiotic stresses.

Claim 144; SEQ ID NO 2212; 577pp + Sequence Listing; English.

The invention relates to identifying a stress condition to which a plant cell has been exposed, comprising: (a) contacting nucleic acid representative of expressed polynucleotides in the plant cell with an array or probes representative of the plant cell genome; and (b) detecting a profile of expressed polynucleotides in the plant cell characteristic of a stress response. The method is useful in the production of transgenic plants, cells and seeds and in producing plants with increased tolerance to abiotic stress. The present sequence is that of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used in methods of the invention. Note: The sequence data for this patent is

CC not represented in the printed specification but is based on sequence  
CC information supplied to Derwent by the European Patent Office  
XX  
SQ Sequence 3387 BP; 870 A; 781 C; 654 G; 1082 T; 0 U; 0 Other;  
Query Match 89.1%; Score 1152; DB 6; Length 3387;  
Best Local Similarity 96.4%; Pred. No. 0;  
Matches 1207; Conservative 0; Mismatches 0; Indels 45; Gaps 1;  
QY 22 ATGGAATAACTCAAGTAATCTATGTCAATGGTGAAGACATCATCTCGTAGTGAGAT 81  
DB 1 ATGGAATAACTCAAGTAATCTATGTCAATGGTGAAGACATCATCTCGTAGTGAGAT 60  
QY 82 TACTTGATCTTTTCCTCAACCTGTTGTTTTCAGATGTTTTCAGATCATAACAACATAGAA 141  
DB 61 TACTTGATCTTTTCCTCAACCTGTTGTTTTCAGATGTTTTCAGATCATAACAACATAGAA 120  
QY 142 GCAAACTAACAGGTGAGGAGGAATGAATACCAACATGGAATCGAATCTTCGTCGTA 201  
DB 121 GCAAACTAACAGGTGAGGAGGAATGAATACCAACATGGAATCGAATCTTCGTCGTA 180  
QY 202 CCTCAATCGAAGAACTATAGAAATGGCTTCGTAATTTCCATTTACGTTGTTCTTTGTCCTT 261  
DB 181 CCTCAATCGAAGAACTATAGAAATGGCTTCGTAATTTCCATTTACGTTGTTCTTTGTCCTT 227  
QY 262 GCTTGCNAAGCACTTTCTACAATTTTGGCAGAGTTTACTATGAAATGGTGGAGAGT 321  
DB 228 -----AGTTTACTATGAAATGGTGGAGAGT 255  
QY 322 ACATGATGGAAACACTGTTCGAACCTAAATCGGCTTCCTGTTCTGTTCTCTTCGCTTC 381  
DB 256 ACATGATGGAAACACTGTTCGAACCTAAATCGGCTTCCTGTTCTGTTCTCTTCGCTTC 315  
QY 382 TTTTCCAAACCAAAATCCCAACCAACAGAGAGATTTCAGAAAGTTCTTCCTTC 441  
DB 316 TTTTCCAAACCAAAATCCCAACCAACAGAGAGATTTCAGAAAGTTCTTCCTTC 375  
QY 442 ACCATTCTTGGATCAGTTTACATCGTTACTGGACTATTAGTGTCTGCTAACTCTTATATG 501  
DB 376 ACCATTCTTGGATCAGTTTACATCGTTACTGGACTATTAGTGTCTGCTAACTCTTATATG 435  
QY 502 TCTCTGTTGGTTACTACTACTACAGTTTCTATCTTCTCCCTCAATCTGGCCTCACAA 561  
DB 436 TCTCTGTTGGTTACTACTACTACAGTTTCTATCTTCTCCCTCAATCTGGCCTCACAA 495  
QY 562 TTGGCCTTCACGCTCTTCTCATATTTCTAACTCGCAGAGTTTCACACTTCATT 621  
DB 496 TTGGCCTTCACGCTCTTCTCATATTTCTAACTCGCAGAGTTTCACACTTCATT 555  
QY 622 GTGAATCTCTGTTCTCTTACTATTCTCTGCTCTCTGCTCAACACTGATTTCG 681  
DB 556 GTGAATCTCTGTTCTCTTACTATTCTCTGCTCTCTGCTCAACACTGATTTCG 615  
QY 682 GAAACACAGCAAAAGTGTCTAGAGPAAATATGTGATAGGATATATGTACCAATTGGT 741  
DB 616 GAAACACAGCAAAAGTGTCTAGAGPAAATATGTGATAGGATATATGTACCAATTGGT 675  
QY 742 GCTTCTGCTGGATTGGATTGCTGCTATCCCTGCTACACTGATCTCAGGAGTTTGA 801  
DB 676 GCTTCTGCTGGATTGGATTGCTGCTATCCCTGGTACAACTGATCTCAGGAGTTTGA 735  
QY 802 AAGAAGCAAACTTCTCAACGGTCACTGACTTGGTGGCTTTACCAATCTCTAGTTGCAAGC 861  
DB 736 AAGAAGCAAACTTCTCAACGGTCACTGACTTGGTGGCTTTACCAATCTCTAGTTGCAAGC 795  
QY 862 TGTGTGTTCTCATAGACTTTTCGACGGGGAGTGAAACTTTTAAACAGTGCAGATG 921  
DB 796 TGTGTGTTCTCATAGACTTTTCGACGGGGAGTGAAACTTTTAAACAGTGCAGATG 855  
QY 922 GAAACTACAACTGGGAAAGTGCATAGCTTATGACTTTGGCTCGATAGCTATTTC 981  
DB 856 GAAACTACAACTGGGAAAGTGCATAGCTTATGACTTTGGCTCGATAGCTATTTC 915

QY 982 TGGCAAGTCTACACCAATTCGGCTCGTGGAGTATCTTTGAGTCACTTCTGTGTTCTCC 1041  
DB 916 TGGCAAGTCTACACCAATTCGGCTCGTGGAGTATCTTTGAGTCACTTCTGTGTTCTCC 975  
QY 1042 AATTCCATAACTGCTGTGGATTGCTATAGTTCCAGTTGTAGCAGTGATTTGTTTCCAT 1101  
DB 976 AATTCCATAACTGCTGTGGATTGCTATAGTTCCAGTTGTAGCAGTGATTTGTTTCCAT 1035  
QY 1102 GATAAAATGAACGCGTCAAGATCTTCTCCATCATTTTGTAGTATCTTGGGGATTTCATTCA 1161  
DB 1036 GATAAAATGAACGCGTCAAGATCTTCTCCATCATTTTGTAGTATCTTGGGGATTTCATTCA 1095  
QY 1162 TTTGTCTATCAGCACTACTCTCGAAGAAAGTTTGAAGACTAGCCACACAAGTCTGTGA 1221  
DB 1096 TTTGTCTATCAGCACTACTCTCGAAGAAAGTTTGAAGACTAGCCACACAAGTCTGTGA 1155  
QY 1222 GGAGATCTCTACTACTACTCTGCTGAGGAAGTTCACAAAACATACATAGTG 1273  
DB 1156 GGAGATCTCTACTACTACTCTGCTGAGGAAGTTCACAAAACATACATAGTG 1207  
RESULT 3  
ADA68246  
ID ADA68246 standard; DNA; 3387 BP.  
XX  
AC ADA68246;  
XX  
DT 20-NOV-2003 (first entry)  
XX  
DE Arabidopsis thaliana gene, SEQ ID 908.  
XX  
KW Plant; bacterial infection; fungal infection; viral infection; da.  
XX  
OS Arabidopsis thaliana.  
XX  
PN WO2003000898-A1.  
XX  
PD 03-JAN-2003.  
XX  
PF 22-JUN-2001; 2001WO-IB001105.  
XX  
PR 22-JUN-2001; 2001WO-IB001105.  
XX  
PA (SYGN ) SYNGENTA PARTICIPATIONS AG.  
XX  
PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;  
PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;  
XX  
WPI; 2003-175290/17.  
DR  
XX  
CC Identifying at least one gene involved in plant resistance or response to  
pathogenic infection for conferring resistance or tolerance to a plant to  
bacterial, fungal or viral infection by determining or detecting plant  
gene expression.  
PS Claim 6; SEQ ID NO 908; 899pp; English.  
XX  
CC The present invention relates to a method (M1) for identifying genes  
involved in plant resistance or response to pathogenic infection. M1  
comprises identifying a gene whose expression is significantly altered in  
the incompatible interaction of plant gene expression relative to  
CC expression of the gene in an uninfected plant, in a mutant plant that  
does not express a gene associated with response to pathogenic infection,  
or in a corresponding incompatible or compatible interaction. (M1) is  
useful for conferring resistance to resistance or tolerance to a plant to  
bacterial, fungal or viral infection. The present sequence was used to  
illustrate the invention.  
XX  
SQ Sequence 3387 BP; 870 A; 781 C; 654 G; 1082 T; 0 U; 0 Other;  
Query Match 89.1%; Score 1152; DB 6; Length 3387;  
Best Local Similarity 96.4%; Pred. No. 0;  
Matches 1207; Conservative 0; Mismatches 0; Indels 45; Gaps 1;

```
Qy 22 ATGGAATAAATCAAGTAATCTATGTCAATAGGTAAGCAAGATGCATCTCGTAGTAGAT 81
Db 1 ATGGAATAAATCAAGTAATCTATGTCAATAGGTAAGCAAGATGCATCTCGTAGTAGAT 60
Qy 82 TACTTGATTTCTTTTCGCTAACTGTTGTTTGTGATGTTTTTCAGATCATATAACAATAGAA 141
Db 61 TACTTGATTTCTTTTCGCTAACTGTTGTTTGTGATGTTTTTCAGATCATATAACAATAGAA 120
Qy 142 GCAAACTTAAACAGGTACAGGAGAAATGAATACCAACATGGAATCGAATCTTCCTCGTA 201
Db 121 GCAAACTTAAACAGGTACAGGAGAAATGAATACCAACATGGAATCGAATCTTCCTCGTA 180
Qy 202 CCTCAATCGAAGAACTATAAGAAATGGCTTCGTATTTCCATTTACGTGTTCTTTGTCTTT 261
Db 181 CCTCAATCGAAGAACTATAAGAAATGGCTTCGTATTTCCATTTACGT----- 227
Qy 262 GCTTGCCCAAGCACTTTCTACAAATTTTGGGCAGAGTTTACTATGAAATGGTGGGAAGAT 321
Db 228 -----AGTTTACTATGAATGGTGGGAAGAT 255
Qy 322 ACATGGATGGAAACATTTGTCOAATAATCGGCTTCCCTGTTCTGTTCTTTCCTCGCTTC 381
Db 256 ACATGGATGGAAACATTTGTCOAATAATCGGCTTCCCTGTTCTGTTCTTTCCTCGCTTC 315
Qy 382 TTTTCCCAACCAAAATCCCAACCAACAGAGATTTTCAGAAAGTTCTCTTCCTTC 441
Db 316 TTTTCCCAACCAAAATCCCAACCAACAGAGATTTTCAGAAAGTTCTCTTCCTTC 375
Qy 442 ACCATTTCTGGATCAGTTTACATCGTTACTCGAGCTATTAGTGTCTGCTAACTCTTATATG 501
Db 376 ACCATTTCTGGATCAGTTTACATCGTTACTCGAGCTATTAGTGTCTGCTAACTCTTATATG 435
Qy 502 TCCTCTGTTGGTTTACTATACATTACAGTTTCTACTTTTCTCCCTCACTTTGGCCTCACA 561
Db 436 TCCTCTGTTGGTTTACTATACATTACAGTTTCTACTTTTCTCCCTCATCTTGGCCTCACA 495
Qy 562 TTGGCCTTCACTGCTTCTCTCATATTTTCTAACTCGCAGAGTTTCACACCTTTTCATT 621
Db 496 TTGGCCTTCACTGCTTCTCTCATATTTTCTAACTCGCAGAGTTTCACACCTTTTCATT 555
Qy 622 GTGAATTTCTGTTTCTCTTACTATTTCCTCTGCTGCTCCTCGGTCAACACATGATTCG 681
Db 556 GTGAATTTCTGTTTCTCTTACTATTTCCTCTGCTGCTCCTCGGTCAACACATGATTCG 615
Qy 682 GAAACACAGCAAAAGTGTCTAGAGTAAATATGTGATAGGGATAAATGTACCATTTGGT 741
Db 616 GAAACACAGCAAAAGTGTCTAGAGTAAATATGTGATAGGGATAAATGTACCATTTGGT 675
Qy 742 GCTTCTGCTGGATTGGATTGCTGCTATCTTCCTGCTTACACATGATCCTCAGGAAGTTTA 801
Db 676 GCTTCTGCTGGATTGGATTGCTGCTATCTTCCTGCTTACACATGATCCTCAGGAAGTTTA 735
Qy 802 AAGAACCAACATTTCTCAACGGTCACTGACTTGGTGGCTTACCAATCTCTAGTTGCAAGC 861
Db 736 AAGAACCAACATTTCTCAACGGTCACTGACTTGGTGGCTTACCAATCTCTAGTTGCAAGC 795
Qy 862 TGTGTGTTCTCATAGACTTTTCGACAGGGGAGTGGAACCTTTTAAACAGTGAGATG 921
Db 796 TGTGTGTTCTCATAGACTTTTCGACAGGGGAGTGGAACCTTTTAAACAGTGAGATG 855
Qy 922 GAAACTACAACCTGGGGAAGTGCATACCTTATGACTTTGGCTCGATAGCTATTTC 981
Db 856 GAAACTACAACCTGGGGAAGTGCATACCTTATGACTTTGGCTCGATAGCTATTTC 915
Qy 982 TGGCAAGTCTACACATTTGGGCTGCTGGGACTGATCTTTAGTCACTCTCTGTTCTCC 1041
Db 916 TGGCAAGTCTACACATTTGGGCTGCTGGGACTGATCTTTAGTCACTCTCTGTTCTCC 975
Qy 1042 AATTCATAACTGCTGTGGGATTCCTATAGTTCAGTTGTAGCAGTGAATTTTTCAT 1101
Db 976 AATTCATAACTGCTGTGGGATTCCTATAGTTCAGTTGTAGCAGTGAATTTTTCAT 1035
```

```
Qy 1102 GATAAAATGAACGGTCAAAAGATCTTCTCCATCATTTTAGCTATCTGGGATTCATTTC 1161
Db 1036 GATAAAATGAACGGTCAAAAGATCTTCTCCATCATTTTAGCTATCTGGGATTCATTTC 1095
Qy 1162 TTTGTCTATCAGCACTACCTCGACGAAAGAGTTGAAGACTAGCCACACAAGTCTCTGTA 1221
Db 1096 TTTGTCTATCAGCACTACCTCGACGAAAGAGTTGAAGACTAGCCACACAAGTCTCTGTA 1155
Qy 1222 GGAGATCCTCATCTACTACCTGCTGAGGAAGGTCACAAACATACATAGTG 1273
Db 1156 GGAGATCCTCATCTACTACCTGCTGAGGAAGGTCACAAACATACATAGTG 1207

RESULT 4
AAA97923
ID AAA97923 standard; DNA; 1194 BP.
AC AAA97923;
XX
DT 19-JAN-2001 (first entry)
XX
DE A. thaliana PUP1 DNA #5.
XX
KW PUP1; transgenic plant; nucleobase transporter; apical dominance;
flowering behaviour; senescence; pesticide distribution; ds.
XX
OS Arabidopsis thaliana.
XX
PN DE19907209-A1.
XX
PD 24-AUG-2000.
XX
PF 19-FEB-1999; 99DE-01007209.
XX
PR 19-FEB-1999; 99DE-01007209.
XX
PA (FROM/) FROMMER W.
XX
PI Gillissen B, Buerkle L, Andre B, Frommer WB;
XX
WP1; 2000-566202/53.
XX
DR
XX
PT Nucleic acid, useful for producing transgenic plants with altered
nucleobase transporter, encodes a nucleobase transporter protein of
Arabidopsis thaliana.
XX
PS Claim 1f; Page 14; 24pp; German.
XX
CC This invention describes a novel nucleic acid encoding a plant nucleobase
transporter (I). (I) is produced by complementation of a nucleobase
transport (NBT)-defective host cell with a plant gene bank by selection
of NBT-positive cells. (I) is used to isolate homologous sequences from
bacteria, fungi, plants, animals and humans, for expression of the
encoded protein (II) in prokaryotic or eukaryotic cells, for inhibiting
expression of (II) (when in antisense orientation), and to produce
transgenic crop plants. The transgenic plants have modified nucleobase
transport properties, e.g. altered affinity and substrate specificity
that may result in more efficient nucleobase transport in leaves, changes
in apical dominance, flowering behaviour and senescence, or improved
distribution of pesticides. This sequence encodes the Arabidopsis
thaliana PUP1 protein which is described in the method of the invention
XX
SQ Sequence 1194 BP; 298 A; 279 C; 231 G; 386 T; 0 U; 0 Other;

Query Match 63.2%; Score 816.6; DB 3; Length 1194;
Best Local Similarity 81.8%; Pred. No. 6.1e-233;
Matches 971; Conservative 0; Mismatches 204; Indels 12; Gaps 2;

Qy 107 TGTGTTTGTGTTTTCAGATCATATAACAATAGCAACCTTAACAGGTACAGGAGAA 166
Db 20 TGAATGTTAAATTTTTCAGGTGACACAGAACTTAGAAGCAAACTTTATAGATCATGAGGTGG 79
Qy 167 TGAATACCACCATGGAAATCGAATCTTTCGTCCTGCTCACTCAATCGAAGACTATAAGAAAT 226
```



```
QY 504 CTCTGTTGGTTTACTATATACAGTTTCTACTTTCTCCCTCATCTTGGCCTCACAAATT 563
Db 285 TTCCATTGAGCTGCTTTACCTACCTGTTTCTACCTTTCCCTGATCTGTGCATCACAGTT 344
QY 564 GGCCTTCACTGCTTCTCTCATATTTTCTAAACTCGCAGAAAGTTTCACACCTTTTCATTGT 623
Db 345 AGCCTTACCCTTCTCTTCTTATTTACTCACTCACAAAACCTTACTCTCTATCATTTT 404
QY 624 GAATTTCTGTTTCTCTTACTATTTCTCTGCCCCTCTCTGCTGCTCAACACTGATTCGGA 683
Db 405 GAATTTCTTCTTCTCTCTCACTATATCTTCCACCTCTCTGCAATTTAATAACGAGGAATC 464
QY 684 AAACACACAAAAGTCTAGAGTAAATATGTGATAGGATATATGTACCATTTGTGTC 743
Db 465 AGATTTCCAAAAGAGTTACAAAAGGAGATGTGTCAAAAGGTTTCGTATGCACCGTTGTGTC 524
QY 744 TTCTGCTGGGATTTGGATTGCTCTATCCCTGCTACACTGATCTCTCAGGAAGGTTTAA 803
Db 525 ATCTGTGGGTTTGGTCTACTCTTATCCCTACACAGCTAGCCCTTGTAAAGTTTAA 584
QY 804 GAAGCAAACTTCTCAACGGTCACTGACTTGTGCTGCTTTACCAATCTCTAGTTGCAAGCTG 863
Db 585 GAAGCAAACTTCTCAGAAGTTATAAATATGATAATCTACATGAGTCTAGTGGCCAGTTG 644
QY 864 TGTGTTCTCATAGGACTTTTCGCAAGCGGGAGTGGAAACTTTAACTGAGATGGA 923
Db 645 TGTAGCTGTGGGGCTTTTGTCTAGTAGCGAGTGGAAACTTTTGAGCAGTGAATGGA 704
QY 924 AAATACAAACTGGGGAAGTGCATACGTTATGACTTTTGGCCCTCGATAGTATTTCTCTG 983
Db 705 AAATACAAACTTGGGAAGGTATCTATGTCATGAACCTAGTGTGGACAGCTGTACTCTG 764
QY 984 GCAAGTCTACACCAATGCGCTGCTGGGACTGATCTTTGAGTCACTCTTCTGTGTTCTCAA 1043
Db 765 GCAGTATTTCTCCATCGGTTGCACAGGACTGATCTTTCGAGCTTTCCTCCCTATTCTCAA 824
QY 1044 TTCCATACTCTGTGGGATTCCTATAGTTCCAGTTGTAGCAGTATGTTTCCATGA 1103
Db 825 TGCAATAAGCCCTTTGGGACTCCCGGTTTCTCTATCCTGTGCTGATCATTTTCCATGA 884
QY 1104 TAAATGAACGGTCAAAAGATCTTCTCCATCATTTTATGCTATCTGGGATTCATTTCAAT 1163
Db 885 CAAATGAACGGCTTAAGGTGATTTCTATGATTTAGTATTTGGGTTTCGTATCTTA 944
QY 1164 TGTCTATCAGCACTACCTCGACGAAAGAGTTGAAGACTGACCAACAAGTCTGTGAGG 1223
Db 945 TGTCTACCAACAATATCTTGTATGAACAACTTGAAGAAAAGTAATGAATACCAACAAC 1004
QY 1224 AGATCCTCATCTACTACCTGTGAGGAAGG 1253
Db 1005 AGAATCCCTGACCGACCAAGAGCAGGAAGG 1034
```

## RESULT 6

```
ADG88190
ID ADG88190 standard; cDNA; 1047 BP.
XX
AC
XX
XX
ADG88190;
XX
DT 22-APR-2004 (first entry)
XX
DE A. thaliana RPP4-upregulated pathogen infection-related gene #632.
XX
KW Pathogen infection-related gene; plant; Peronospora parasitica;
KW defence mechanism; RPP4; pathogen resistance; transgenic plant; oomycete;
KW fungus; bacterium; virus; nematode; insect; aphid; gene; ss.
XX
OS Arabidopsis thaliana.
XX
FN WO200222675-A2.
XX
PD 21-MAR-2002.
```

```
XX
PF
XX
PR 14-SEP-2001; 2001WO-US028506.
PR 15-SEP-2000; 2000US-0232778P.
PR 22-JUN-2001; 2001US-0300183P.
XX
XX (SYGN ) SYNGENTA PARTICIPATIONS AG.
PA (LYNC-) UNIV NORTH CAROLINA.
PA (GLAZ/) GLAZEBROOK J.
PA (WANG/) WANG X.
PA (DANG/) DANG J L.
PA (EULG/) EULGEM T.
PA (ZHUT/) ZHU T.
XX
XX Glazebrook J, Wang X, Dangl JL, Eulgem T, Zhu T;
PI WPI; 2002-292409/33.
XX
XX Novel isolated polynucleotide, useful for conveying pathogen resistance
PT to plants, and for identifying plants infected with a pathogen.
XX
XX Claim 3; SEQ ID NO 632; 605pp; English.
PS
XX
XX The invention relates to 691 Arabidopsis thaliana genes (ADG87559--
CC ADG87557) whose expression is altered in response to pathogen infection,
CC and to homologues of these genes from other plants or fungi: especially
CC from maize, soybean, barley, alfalfa, sunflower, canola (oilseed rape),
CC cotton, peanut, sorghum, tobacco, sugarbeet, rice or wheat. The
CC Arabidopsis plants infected with the oomycete Peronospora parasitica,
CC indicating that they play a role in defence mechanisms. The genes of the
CC invention are regulated by RPP7 or RPP8 which act via unconventional
CC signalling cascades, or by the RPP4-dependent pathway. The invention also
CC relates to polypeptides encoded by the pathogen infection-related genes;
CC promoter motifs from pathogen infection-related genes (ADG88243-ADG88327)
CC ; expression cassettes, host cells and pathogen-resistant transgenic
CC plants and their progeny comprising a polynucleotide of the invention;
CC and a method of identifying a plant cell infected with a pathogen. The
CC polynucleotide sequences and methods of the invention are useful for
CC identifying plants infected with a pathogen, and for conferring
CC resistance to pathogens such as oomycetes, fungi, bacteria, viruses,
CC nematodes and insects (e.g., aphids). The present sequence represents an
CC Arabidopsis thaliana gene whose expression is altered in response to
CC Peronospora parasitica infection. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1047 BP; 309 A; 225 C; 199 G; 314 T; 0 U; 0 Other;
```

```
Query Match 31.7%; Score 410; DB 6; Length 1047;
Best Local Similarity 65.0%; Pred. No. 1.7e-111;
Matches 624; Conservative 0; Mismatches 330; Indels 6; Gaps 1;

QY 293 GAGTTTACTATGAAATGGTGGGAAGAGTAGCATGGATGGGAACACTTGTCCAACTAATCG 352
Db 71 GAGTATACTATGACACGAGGAAACAGTAATGGCTAGCAACGGTAGTTCACTTGTG 130
QY 353 GCTTCCCTGTTTCTGTTTCTCTTCGCTTCTTTTCCCAACCAAAATCCCAACCAACAG 412
Db 131 GCTTTCCTGTGCTACTTCCCATATTATATCTGTCAATTTAAACACATGCAACAACATGATA 190
QY 413 AAGCAGATTTTCAGAAAGTTCTCTTCCTCCATTCTTGGATCAGTTTACATCCCTACTG 472
Db 191 GAGATGGAAGAAAGAACCTCACCT-----AGGAACCGTGTATTGGTTTACGTAGTCTG 244
QY 473 GACTATTAGTGTCTGCTAACTCTTATATGTCTCTGTTGGTTTACTACTACTACAGTTT 532
Db 245 GACTTCTTGTAGGACGAGATTGCTATCTGTACTCTCATTCGACTTCTTTACTTACCCGTT 304
QY 533 CTACTTTTCTCCCTCACTTGGCCCTTCAAAATTTGGCCCTTCACTGCTTCTTCTCATATTTC 592
Db 305 CTACCTATTCCCTGATCTGTGCATCTCAGTTAGCCTTCAATGCTTCTTCTTCTTATTTC 364
```





Db 743 TTAGAACCTAGTGGACAGCTGTTACCTGGGAGGTATTCTCCATCGGTGGACAGAC 802  
Qy 1013 TGATCTTTGAGTCATCTTCTGTGTTCTCCAAATCCATAACTGCTGTGGATTGCTATAG 1072  
Db 803 TGATCTTCGAGCTCTCCCTCTCTATTCTCAAAATGAAGGTTTGGGACTCCAGTGG 862  
Qy 1073 TTCAGTTGTAGCAGTGAATGTTTTCATGATGAATAAATGAAGCGTCAAGATCTTCTCCA 1132  
Db 863 TTCCTATCTTGGCTGTAAATCAATTTTCCATGACAAAATGAATGGGTTAAAGGTGATTCTA 922  
Qy 1133 TCATTTTGTAGCTATCTGGGATTCATTTTCTATCTATGCTATGCTACCTGCTGACAAAAGA 1192  
Db 923 TGATCTTAGCTATTTGGGTTTCACTTCTCTATGCTACCAACATATCTTGATGACAAA 982  
Qy 1193 AGTTGAAGACTAGCCACAAAGTCTGTGAGAGATCCTCATCTACTACTGCTGAGGAAG 1252  
Db 983 ACTTGAAGAAAATCATGAATATCAACACAGAAATCCCTGACCCACAGAGCAGAAG 1042

RESULT 8  
ID ABX56692 standard; DNA; 592 BP.  
XX AC ABX56692;  
XX DT 20-FEB-2003 (first entry)  
XX DE Arabidopsis thaliana polynucleotide #44.

XX KW Thale cress; gene; ds; genetic manipulation; plant; biosynthesis;  
XX KW genetic modification; environmental stress; disease resistance;  
XX KW fungicide; insecticide; stress tolerance.  
XX OS Arabidopsis thaliana.

XX PN US2002040489-A1.  
XX PD 04-APR-2002.  
XX PF 26-JAN-2001; 2001US-00770152.  
XX PR 27-JAN-2000; 2000US-0178503P.

XX PA (GORL/) GORLACH J.  
XX PA (ANY/) AN Y.  
XX PA (HAMI/) HAMILTON C M.  
XX PA (PRIC/) PRICE J L.  
XX PA (RAIN/) RAINES T M.  
XX PA (YUY/) YU Y.  
XX PA (RAME/) RAMEAKA J G.  
XX PA (PAGE/) PAGE A.  
XX PA (MATH/) MATHEN A V.  
XX PA (LEDF/) LEDFORD B L.  
XX PA (WOES/) WOESSNER J P.  
XX PA (HAAS/) HAAS W D.  
XX PA (GARC/) GARCIA C A.  
XX PA (KRIC/) KRICKER M.  
XX PA (SLAT/) SLATER T.  
XX PA (DAVI/) DAVIS K R.  
XX PA (ALLE/) ALLEN K.  
XX PA (HOFF/) HOFFMAN N.  
XX PA (HURB/) HURBAN P.

XX Gortlach J, An Y, Hamilton CM, Price JL, Raines TM, Yu Y;  
PI Rameaka JG, Page A, Mathew AV, Ledford BL, Woessner JP, Haas WD;  
PI Garcia CA, Kricker M, Slater T, Davis KR, Allen K, Hoffman N;  
PI Hurban P;  
XX WPI; 2003-110410/10.

XX Novel Arabidopsis thaliana nucleic acid useful for identifying homologous  
PT or related genes, and to create genetically modified and transgenic

PT organisms, such as plant cells and plants.  
XX Claim 1; SEQ ID NO 44; 45pp; English.  
XX The invention relates to Arabidopsis thaliana nucleic acid sequences. The  
CC DNA sequences and the polypeptides they encode are useful for identifying  
CC homologous or related genes, for producing compositions that modulate the  
CC expression or function of the polypeptides, for mapping functional  
CC regions of the protein, in diagnosis, for studying associated  
CC physiological pathways, for genetic manipulation of cells, preferably  
CC plant cells, in screening assays of various plant strains to determine  
CC the strains that are capable of withstanding a particular disease or  
CC environmental stress, for enhancing or inhibiting production of  
CC biosynthetic products in plants and to create genetically modified and  
CC transgenic organisms, such as plant cells and plants. Transgenic plants  
CC are useful for introducing or improving disease resistance and stress  
CC tolerance in plants, screening biologically active agents, such as  
CC fungicides and insecticides, and for elucidating biochemical pathways.  
CC Sequences ABX56649-ABX57647 represent Arabidopsis thaliana  
CC polynucleotides of the invention. Note: The sequence data for this patent  
CC did not form part of the printed specification but was obtained in  
CC electronic format directly from USPTO at seqdata.uspto.gov/sequence.html  
XX SQ Sequence 592 BP; 135 A; 158 C; 96 G; 203 T; 0 U; 0 Other;

Query Match 27.4%; Score 354.4; DB 10; Length 592;  
Best Local Similarity 81.8%; Pred. No. 5.3e-95;  
Matches 409; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

Qy 180 GGAAATCGAATCTTCGTCCTGCTCAATCAAGAACTATAAGAAATGGCTTCGTATTTC 239  
Db 93 GGAACCTGATCATTTTCAGTACTCTCAACGAAGAACTGTAAGAGGTGGCTCCGTGCTC 152  
Qy 240 CATTTACGTTCTTTGTCCTTCTGCTGCAAGCACTTTCTACAAATTTTGGGCGAGTTTA 299  
Db 153 CATATACGCAATCTTTGTCTATCTTCTGCCAACCACTTGCTACAGTTCTGGGTAGACTGA 212  
Qy 300 CTATGAATGTTGGGAGAGTACATGATGGGAGCACTTGTCACAACTAATCGGCTTCCC 359  
Db 213 CTATGAATGTTGGGAGGAAAGCACATATGTTGTAACACTTCTTCAACTCATTTGGCTTCCC 272  
Qy 360 TGTTCGTTCTCTTCGCTTCTTTTCCAAACCAAAATCCCAAAACCAACAGACAGA 419  
Db 273 TGTACTGATTTCTCTTCGCTTCTTTCTCGAATCAGGCAACCAAAATCAACAGATACAA 332  
Qy 420 TTTCAGAAAGTTCTCTTCCTTCCACTTCTTGATCAGTTTACATCGTTACTGGACTATT 479  
Db 333 TTTTCAGTCAGTCCCTTCTCTTCCACACCTTGCATCGTTTACTTGTGCATCGGACTGCT 392  
Qy 480 AGTGTCTGCTAACTCTTATATGCTCTGTTGTTTACTATATCTTACAGTTTCTACTTTT 539  
Db 393 AGTGTCTGCTTATGCTTATTTGCTGCGATGTTGGTGTCTTACTTACAGTCTCTACTTT 452  
Qy 540 CTCCTCATCTTGGCCTTCACAAATTTGGCCTTCACTGCTTCTTCTCATATTTTCTAAACTC 599  
Db 453 CTCCTCATCTTGGCCTTCACAGTTGGCCTTCACTGCTTCTTCTCATATTTTCTTAACTC 512  
Qy 600 GCAGAAAGTTACACCTTTTCATTTGTAATTTCTGTTTCTCTTACTACTATTTCTCTGCTC 659  
Db 513 GCAAAAGTTCACTCTTGTATGATGATCAATTTCTTGTCTCTCTTACGTTTCTCTGCTC 572  
Qy 660 CCTCGTGGTCAACACTGATT 679  
Db 573 CCTCGTGGTCAACACTGATT 592

RESULT 9  
ID AAA97921 standard; DNA; 1145 BP.  
XX AC AAA97921;  
XX DT 19-JAN-2001 (first entry)

XX A. thaliana PUP1 DNA #3.  
XX PUP1; transgenic plant; nucleobase transporter; apical dominance;  
KW flowering behaviour; senescence; pesticide distribution; ds.  
XX Arabidopsis thaliana.  
OS  
XX DE19907209-A1.  
XX 24-AUG-2000.  
XX 19-FEB-1999; 99DE-01007209.  
XX 19-FEB-1999; 99DE-01007209.  
XX (FROM/) FROMMER W.  
XX Gillissen B, Buerkle L, Andre B, Frommer WB;  
PI WPI; 2000-566202/53.  
XX Nucleic acid, useful for producing transgenic plants with altered  
PT nucleobase transport, encodes a nucleobase transporter protein of  
PT Arabidopsis thaliana.  
XX  
PS Claim 1f; Page 12-13; 24pp; German.  
XX This invention describes a novel nucleic acid encoding a plant nucleobase  
CC transporter (I). (I) is produced by complementation of a nucleobase  
CC transport (NBT)-defective host cell with a plant gene bank by selection  
CC of NBT-positive cells. (I) is used to isolate homologous sequences from  
CC bacteria, fungi, plants, animals and humans, for expression of the  
CC encoded protein (II) in prokaryotic or eukaryotic cells, for inhibiting  
CC expression of (II) (when in antisense orientation), and to produce  
CC transgenic crop plants. The transgenic plants have modified nucleobase  
CC transport properties, e.g. altered affinity and substrate specificity  
CC that may result in more efficient nucleobase transport in leaves, changes  
CC in apical dominance, flowering behaviour and senescence, or improved  
CC distribution of pesticides. This sequence encodes the Arabidopsis  
CC thaliana PUP1 protein which is described in the method of the invention  
XX  
SQ Sequence 1145 BP; 309 A; 234 C; 226 G; 376 T; 0 U; 0 Other;  
Query Match 22.8%; Score 294.4; DB 3; Length 1145;  
Best Local Similarity 56.3%; Pred. No. 6.2e-77;  
Matches 578; Conservative 0; Mismatches 436; Indels 12; Gaps 1;  
QY 181 GAAATCGAATCTTCGTCCTACTCAATCGAAGAACTATAAGAAATGGCTTCGTATTTCC 240  
DB 28 GAAGGGAATTTTCAACAGAGAGAGAGAGTCAAGTACTCTTGGAGGTTAAGAGTGCT 87  
QY 241 ATTTAGGTGTTCTTGTCTTCTTGGCCAGCACTTTCTACAAATTTGGGCGAGAGTTTAC 300  
DB 88 CTCATGTCACTCTCTCTTCTTCTGAGAGACAATAGCCACTCTCTTAGGTAGACTTTAC 147  
QY 301 TATGAAATGGTGGGAAGTAGTACATGATGGGAACACTTGTCCAACTAATCGCTTCCT 360  
DB 148 TAGGAAAGGGGTAAGACACATGGCTCGAAACCTTGGTTGAGTGGGTTTCT 207  
QY 361 GTTCTGTTTCTCTTCGCTTCTTTTCCAAACCAAAATCCCAACCAACAGAGAGAT 420  
DB 208 TTAACCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 267  
QY 421 TTCAGAAATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 480  
DB 268 AAAAATACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 327  
QY 481 GTGCTGCTAACTCTTATATGTCCTTGTGTTTCTTCTTCTTCTTCTTCTTCTTCTTCT 540  
DB 328 GTTGTGGACATGATTTTGTGACATCAATTTGGGCTTCTTCTTCTTCTTCTTCTTCTTCT 387  
QY 541 TCCCTCATCTGGGCTCAACAATTTGGCCTTCACTGCTTCTTCTTCTTCTTCTTCTTCTTCT 600

DB 388 TCTTTGATCTCTGCGTCGCAATTGGCTTTTAACGCGCTCTTCTCTTACTTCTTAAACTCA 447  
QY 601 CAGAAAGTTACACACCTTTTCTTCTGTAATCTCTGTTTCTTCTTCTTCTTCTTCTTCTTCT 660  
DB 448 CAAAAAATCACACCAATTTATATCTCAATTCATCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 507  
QY 661 CTCGTGGTCAACACTGATTTCGGAA-----AACACAGCAAAAGTGTCTAGAGTA 708  
DB 508 CTTGTTATCCACATGAACCAAGATCTCCCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 567  
QY 709 AAATATGTATAGGATTAATATGTACCAATTTGTTGTTCTTCTTCTTCTTCTTCTTCTTCTTCT 768  
DB 568 AAGTATGTATGGATACATCTGCGCGTCTGCTAGCTCAGCTGTTTCTTCTTCTTCTTCTTCT 627  
QY 769 TCCCTGGTACACTGATCTCTCAGGAAGGTTTAAAGAGCAAACTTCTCAACGGTCACT 828  
DB 628 TCTTTAAGAGATTACCGGTTGAAAGATTTTAAAGAAATACACATTTCAAGGCTATTTTA 687  
QY 829 GACTTGGTGGCTTACCAATCTCTAGTTCGCAAGCTGTGTTCTTCTTCTTCTTCTTCTTCTTCT 888  
DB 688 GACATGGCCACATATCCGCTCTATGTTAGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 747  
QY 889 AGCGGGAGTGGAAACCTTTTAAAGTGAATGGAATACTACAACTGGGGAAGTGCCA 948  
DB 748 AGTGTGGTGGAAAAAGCTGAGTACAGAAATGGAAGAGTTTCAACTAGGGAAGAGCTCA 807  
QY 949 TACGTTATGACTTTGGCCTCGATAGCTATTTCTTCTGCGCAAGTCTACACCATTTGGCGTCGTG 1008  
DB 808 TACATTTTGTATAAACAATCCGTTCAACAGATATCATGGCAAGCTTTGTTGATGGAAGTGT 867  
QY 1009 GCACTGATCTTTGAGTCACTTCTGTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1068  
DB 868 GGTGTGATATCGAAGTTTCACTCGCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 927  
QY 1069 ATAGTTCAGTTGTAGCAGTGAATTTTCCATGATAAATAAGACGCTCAAGATCTTCTTC 1128  
DB 928 GTTGTGCTGTTCTTCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 987  
QY 1129 TCCATCATTTTAGCTATCTGGGATTCATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1188  
DB 988 GCAATGTTTGTGGCAATCTGGGATTTGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1047  
QY 1189 AAGAAAG 1194  
DB 1048 AGAAAG 1053  
RESULT 10  
AAC51119  
ID AAC51119 standard; DNA; 1292 BP.  
XX AAC51119;  
AC AAC51119;  
XX  
DT 18-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 67344.  
XX  
KW Hybridisation assay; genetic mapping; gene expression control;  
KW protein identification; signal transduction pathway; metabolic pathway;  
KW promoter; termination sequence; ss.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-00301439.  
XX  
PR 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.

PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131443P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136382P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139432P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149829P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.

PR	07-OCT-1999;	99US-0158029P.
PR	08-OCT-1999;	99US-0158232P.
PR	12-OCT-1999;	99US-0158369P.
PR	13-OCT-1999;	99US-0159293P.
PR	13-OCT-1999;	99US-0159294P.
PR	13-OCT-1999;	99US-0159295P.
PR	14-OCT-1999;	99US-0159329P.
PR	14-OCT-1999;	99US-0159330P.
PR	14-OCT-1999;	99US-0159331P.
PR	14-OCT-1999;	99US-0159637P.
PR	14-OCT-1999;	99US-0159638P.
PR	18-OCT-1999;	99US-0159584P.
PR	21-OCT-1999;	99US-0160741P.
PR	21-OCT-1999;	99US-0160767P.
PR	21-OCT-1999;	99US-0160768P.
PR	21-OCT-1999;	99US-0160770P.
PR	21-OCT-1999;	99US-0160814P.
PR	21-OCT-1999;	99US-0160815P.
PR	22-OCT-1999;	99US-0160800P.
PR	22-OCT-1999;	99US-0160981P.
PR	22-OCT-1999;	99US-0160989P.
PR	25-OCT-1999;	99US-0161404P.
PR	25-OCT-1999;	99US-0161405P.
PR	25-OCT-1999;	99US-0161406P.
PR	26-OCT-1999;	99US-0161359P.
PR	26-OCT-1999;	99US-0161360P.
PR	26-OCT-1999;	99US-0161361P.
PR	28-OCT-1999;	99US-0161920P.
PR	28-OCT-1999;	99US-0161922P.
PR	28-OCT-1999;	99US-0161933P.
PR	23-OCT-1999;	99US-0162142P.

Query Match 22.8%; Score 294.4; DB 3; Length 1292;  
Best Local Similarity 55.6%; Pred. No. 6.6e-77;  
Matches 593; Conservative 0; Mismatches 461; Indels 12; Gaps 1;

QY	141	AGCAAACTTACAGGTCAGAGGAAATGAATACACCATGGGAAATCGAATCTTCGTCCT	200
DB	39	AGAACTCAAGAACTCCATCTTCACGTAATGCGAACTGGAAGGAAATTTTCAACAGA	98
QY	201	ACCTCAATCGAAGCACTATAAGAAATGGCTTCCTATTTCCATTTAGTCTTCTTCTCT	260
DB	99	AGAGAGAGTCACAGTACTCTTGGAGGTTAAGAGTGTCTCTATGCACTCTCTCTT	158
QY	261	TGCTTGCCAGCACTTTTACAAATTTTGGCGAGTTTACTATGAAATGGTGGGAAGAG	320
DB	159	AGCTGGAGAGCAATAGCCACTCTCTTAGTAGACTTTACTAGAAAGGCGTAAAG	218
QY	321	TACATGGATGGGAACACTTGTGCAACTAATCGGCTTCCCTGTCTCTTCTCTCCGCTT	380
DB	219	CACATGGCTCGAAACCTTGGTTTCAGCTTGTAGGTTTCTCTTAAACCTTCTTGTATTA	278
QY	381	CTTTTCCCAACCAAAATCCCAACCAACAGAGCAGATTTTCAGAAAGTCTCTCTCTT	440
DB	279	TTACTTAAAGCTTGAGCGGTCCAGACTAAACCAATTAACCAAACTACTTCTCTCTT	338
QY	441	CACCAATTTTGGATCAGTTTACATCGTTACTGACTATTAGTCTGTCAACTCTTATAT	500
DB	339	CTTGACACTATCTTTAGTGTATATTGCACTTGGCTTGTGTGGACATGATTTT	398
QY	501	GTCTCTGTGGTTTACTATACATACAGTTTCTACTTCTCCCTCATCTTGGCTTCA	560
DB	399	GTACTCATTTGGGCTACTTTACCTTCTGTCTCAACTTTTCTTGTATCTCTGCTCGCA	458
QY	561	ATTGGCTTCACTGCTTCTTCTCATATTTCTAAACTGCGAGAAGTTTCACACTTTAT	620
DB	459	ATTGGCTTTTAAAGCGCTTCTTCTTACTTCTTAAACTCACAAAAATCACACATTTAT	518
QY	621	TGTGAATTTCTGTTTCTCTTACTATTCTCTGCTCCCTCTGCTGCTCAACACTGATTC	680
DB	519	ACTCAATTCATTTGTTCTCTTAAACCAATGCTCTACACTTCTTGTATCAACATGAAC	578
QY	681	GGAA-----AACACAGCAAAAGTGTCTAGAGTAAATATGTGATAGGGATAAT	728

DB	579	AGAACTCTCCTCTTCTTACTTCAAAGTCCGACCAAGTCAAGTATGTGATGATACAT	638
QY	729	ATGTACCAATTTGCTGCTGCTGGGATTTGATTTGCTGCTATCCCTGGTACAACTGATCCT	788
DB	639	CTGCGCGGTGCGTAGCTCAGCTGGTTATTTCTCTGCTGCTTTCTTTAACAGATTACCGGT	698
QY	789	CAGGAAGGTTTTAAAGAAACAAATTTCTCAACGGTCACTGACTTGGTGGCTTTACCAATC	848
DB	699	CGAAAAGATTCTAAAGAAATACACATTTCAAGGCTATTTTAGACATGCGCCACATATCCGT	758
QY	849	TCTAGTTGCAAGCTGTGCTGCTCATAGGACTTTTTCGCAAGCGGGAGTGGAACCTTT	908
DB	759	TATGGTAGCTACTTGTAGTTGCTAGGACTTTTTCGAAAGTGGTGGTGGGAAAAGCT	818
QY	909	AACAAGTGAGATGGAACAACTCAAACTGGGAAAGTCCATACGTTATGACTTTTGGCCTC	968
DB	819	GAGTACAGAAATGGAAGAGTTTCAACTAGGAAAGAGCTCATACATTTTATAAACATCGG	878
QY	969	GATAGCTATTTCTGGCAAGTCTACACCAATTTGGCGTCTGGGAGCTGATCTTTGAGTCATC	1028
DB	879	TTCAACGATATCATGCAAGCTTTGTTGATTGGAAGTGTGGTTGATTATCGAAGTTTC	938
QY	1029	TTCTGTGTTCTCAATTCATTAACCTGCTGTGGGATTCCTATAGTTCAGTTGTAGCAGT	1088
DB	939	ATCGCTTTTTCCTCAATGCTAAGCACTCTTTGTTTACCAGTTGTGCTGTTCTTCTGT	998
QY	1089	GATTGTTTTTCCATGATAAATGCAACGCGTCAAGAGTCTTCTCCATCATTTTAGCTATCTG	1148
DB	999	TGCTCTTCTCCGTGATGAGTGAAGTGAATCAAGTTGGTTCGAATGTTTGGCCATCTG	1058
QY	1149	GGGATTTCAATTTCTATTTGCTATCAGCACTACCTCGACGAAAAAGAAG	1194
DB	1059	GGGATTTGTTCTTCTATGTTATCAGCAATTTATGTCATGATAGAAG	1104
RESULT 11			
AAC37289			
ID	AAC37289 standard; DNA; 1295 BP.		
XX	AAC37289;		
AC	AAC37289;		
XX	17-OCT-2000 (first entry)		
XX	Arabidopsis thaliana DNA fragment SEQ ID NO: 16846.		
DE	Hybridisation assay; genetic mapping; gene expression control;		
KW	protein identification; signal transduction pathway; metabolic pathway;		
KW	promoter; termination sequence; ss.		
XX	Arabidopsis thaliana.		
XX	EP1033405-A2.		
XX	06-SEP-2000.		
XX	25-FEB-2000; 2000EP-00301439.		
XX	25-FEB-1999; 99US-0121825P.		
PR	05-MAR-1999; 99US-01213180P.		
PR	09-MAR-1999; 99US-0123548P.		
PR	23-MAR-1999; 99US-0125788P.		
PR	25-MAR-1999; 99US-0126264P.		
PR	29-MAR-1999; 99US-0126785P.		
PR	01-APR-1999; 99US-0127462P.		
PR	06-APR-1999; 99US-0128234P.		
PR	08-APR-1999; 99US-0128714P.		
PR	16-APR-1999; 99US-0129845P.		
PR	19-APR-1999; 99US-0130077P.		
PR	21-APR-1999; 99US-0130449P.		
PR	23-APR-1999; 99US-0130510P.		
PR	23-APR-1999; 99US-0130891P.		
PR	28-APR-1999; 99US-0131449P.		

PR	30-APR-1999;	99US-0132048P.
PR	30-APR-1999;	99US-0132407P.
PR	04-MAY-1999;	99US-0132484P.
PR	05-MAY-1999;	99US-0132485P.
PR	06-MAY-1999;	99US-0132486P.
PR	07-MAY-1999;	99US-0132487P.
PR	11-MAY-1999;	99US-0132863P.
PR	14-MAY-1999;	99US-0134256P.
PR	14-MAY-1999;	99US-0134218P.
PR	14-MAY-1999;	99US-0134219P.
PR	14-MAY-1999;	99US-0134221P.
PR	18-MAY-1999;	99US-0134370P.
PR	18-MAY-1999;	99US-0134768P.
PR	19-MAY-1999;	99US-0134941P.
PR	20-MAY-1999;	99US-0135124P.
PR	21-MAY-1999;	99US-0135353P.
PR	24-MAY-1999;	99US-0135622P.
PR	25-MAY-1999;	99US-0136021P.
PR	27-MAY-1999;	99US-0136392P.
PR	28-MAY-1999;	99US-0136782P.
PR	01-JUN-1999;	99US-0137222P.
PR	03-JUN-1999;	99US-0137528P.
PR	04-JUN-1999;	99US-0137502P.
PR	07-JUN-1999;	99US-0137724P.
PR	08-JUN-1999;	99US-0138094P.
PR	10-JUN-1999;	99US-0138540P.
PR	10-JUN-1999;	99US-0138884P.
PR	14-JUN-1999;	99US-0139119P.
PR	16-JUN-1999;	99US-0139452P.
PR	16-JUN-1999;	99US-0139453P.
PR	17-JUN-1999;	99US-0139492P.
PR	18-JUN-1999;	99US-0139454P.
PR	18-JUN-1999;	99US-0139455P.
PR	18-JUN-1999;	99US-0139456P.
PR	18-JUN-1999;	99US-0139457P.
PR	18-JUN-1999;	99US-0139458P.
PR	18-JUN-1999;	99US-0139459P.
PR	18-JUN-1999;	99US-0139460P.
PR	18-JUN-1999;	99US-0139461P.
PR	18-JUN-1999;	99US-0139462P.
PR	18-JUN-1999;	99US-0139463P.
PR	18-JUN-1999;	99US-0139750P.
PR	18-JUN-1999;	99US-0139763P.
PR	21-JUN-1999;	99US-0139817P.
PR	22-JUN-1999;	99US-0139899P.
PR	23-JUN-1999;	99US-0140353P.
PR	23-JUN-1999;	99US-0140354P.
PR	24-JUN-1999;	99US-0140623P.
PR	28-JUN-1999;	99US-0140822P.
PR	29-JUN-1999;	99US-0140991P.
PR	30-JUN-1999;	99US-0141287P.
PR	01-JUL-1999;	99US-0141842P.
PR	01-JUL-1999;	99US-0142154P.
PR	02-JUL-1999;	99US-0142055P.
PR	06-JUL-1999;	99US-0142390P.
PR	08-JUL-1999;	99US-0142803P.
PR	09-JUL-1999;	99US-0142920P.
PR	12-JUL-1999;	99US-0142977P.
PR	13-JUL-1999;	99US-0143542P.
PR	14-JUL-1999;	99US-0143624P.
PR	15-JUL-1999;	99US-0144005P.
PR	16-JUL-1999;	99US-0144085P.
PR	16-JUL-1999;	99US-0144086P.
PR	19-JUL-1999;	99US-0144325P.
PR	19-JUL-1999;	99US-0144331P.
PR	19-JUL-1999;	99US-0144332P.
PR	19-JUL-1999;	99US-0144333P.
PR	19-JUL-1999;	99US-0144334P.
PR	19-JUL-1999;	99US-0144335P.
PR	20-JUL-1999;	99US-0144352P.
PR	20-JUL-1999;	99US-0144632P.
PR	20-JUL-1999;	99US-0144884P.
PR	21-JUL-1999;	99US-0144814P.
PR	21-JUL-1999;	99US-0145086P.
PR	21-JUL-1999;	99US-0145088P.
PR	22-JUL-1999;	99US-0145085P.
PR	22-JUL-1999;	99US-0145087P.
PR	22-JUL-1999;	99US-0145089P.
PR	22-JUL-1999;	99US-0145192P.
PR	23-JUL-1999;	99US-0145145P.
PR	23-JUL-1999;	99US-0145218P.
PR	23-JUL-1999;	99US-0145244P.
PR	26-JUL-1999;	99US-0145276P.
PR	27-JUL-1999;	99US-0145913P.
PR	27-JUL-1999;	99US-0145918P.
PR	27-JUL-1999;	99US-0145919P.
PR	28-JUL-1999;	99US-0145951P.
PR	02-AUG-1999;	99US-0146386P.
PR	02-AUG-1999;	99US-0146388P.
PR	02-AUG-1999;	99US-0146389P.
PR	03-AUG-1999;	99US-0147038P.
PR	04-AUG-1999;	99US-0147204P.
PR	04-AUG-1999;	99US-0147302P.
PR	05-AUG-1999;	99US-0147192P.
PR	05-AUG-1999;	99US-0147260P.
PR	06-AUG-1999;	99US-0147303P.
PR	06-AUG-1999;	99US-0147416P.
PR	09-AUG-1999;	99US-0147493P.
PR	09-AUG-1999;	99US-0147935P.
PR	10-AUG-1999;	99US-0148171P.
PR	11-AUG-1999;	99US-0148319P.
PR	12-AUG-1999;	99US-0148341P.
PR	13-AUG-1999;	99US-0148565P.
PR	13-AUG-1999;	99US-0148684P.
PR	16-AUG-1999;	99US-0149368P.
PR	17-AUG-1999;	99US-0149175P.
PR	18-AUG-1999;	99US-0149426P.
PR	20-AUG-1999;	99US-0149722P.
PR	20-AUG-1999;	99US-0149723P.
PR	20-AUG-1999;	99US-0149929P.
PR	23-AUG-1999;	99US-0149902P.
PR	23-AUG-1999;	99US-0149930P.
PR	25-AUG-1999;	99US-0150566P.
PR	26-AUG-1999;	99US-0150884P.
PR	27-AUG-1999;	99US-0151065P.
PR	27-AUG-1999;	99US-0151066P.
PR	30-AUG-1999;	99US-0151080P.
PR	30-AUG-1999;	99US-0151303P.
PR	31-AUG-1999;	99US-0151438P.
PR	01-SEP-1999;	99US-0151930P.
PR	07-SEP-1999;	99US-0152363P.
PR	10-SEP-1999;	99US-0153070P.
PR	13-SEP-1999;	99US-0153758P.
PR	15-SEP-1999;	99US-0154018P.
PR	16-SEP-1999;	99US-0154039P.
PR	20-SEP-1999;	99US-0154779P.
PR	22-SEP-1999;	99US-0155139P.
PR	23-SEP-1999;	99US-0155486P.
PR	24-SEP-1999;	99US-0155659P.
PR	28-SEP-1999;	99US-0156458P.
PR	29-SEP-1999;	99US-0156596P.
PR	04-OCT-1999;	99US-0157117P.
PR	06-OCT-1999;	99US-0157865P.
PR	07-OCT-1999;	99US-0158029P.
PR	08-OCT-1999;	99US-0158232P.
PR	12-OCT-1999;	99US-0158369P.
PR	13-OCT-1999;	99US-0159293P.
PR	13-OCT-1999;	99US-0159294P.
PR	13-OCT-1999;	99US-0159295P.
PR	14-OCT-1999;	99US-0159329P.
PR	14-OCT-1999;	99US-0159330P.
PR	14-OCT-1999;	99US-0159331P.
PR	14-OCT-1999;	99US-0159637P.
PR	14-OCT-1999;	99US-0159638P.
PR	18-OCT-1999;	99US-0159584P.

PR	21-OCT-1999;	99US-0160741P.	
PR	21-OCT-1999;	99US-0160767P.	
PR	21-OCT-1999;	99US-0160768P.	
PR	21-OCT-1999;	99US-0160770P.	
PR	21-OCT-1999;	99US-0160814P.	
PR	21-OCT-1999;	99US-0160815P.	
PR	22-OCT-1999;	99US-0160980P.	
PR	22-OCT-1999;	99US-0160981P.	
PR	22-OCT-1999;	99US-0160989P.	
PR	25-OCT-1999;	99US-0161404P.	
PR	25-OCT-1999;	99US-0161405P.	
PR	25-OCT-1999;	99US-0161406P.	
PR	26-OCT-1999;	99US-0161359P.	
PR	26-OCT-1999;	99US-0161360P.	
PR	26-OCT-1999;	99US-0161361P.	
PR	28-OCT-1999;	99US-0161920P.	
PR	28-OCT-1999;	99US-0161992P.	
PR	28-OCT-1999;	99US-0161993P.	
PR	29-OCT-1999;	99US-0162142P.	
Query Match 22.8%; Score 294.4; DB 3; Length 1295;			
Best Local Similarity 55.6%; Pred. No. 6.6e-77;			
Matches 593; Conservative 0; Mismatches 461; Indels 12; Gaps 1;			
QY	141	AGCAAACTTAACAGTGCAGAGAGAAATGAATACCACTGGAATCGAATCTTGGTCCGT	200
DB	40	AGAAACTCAAGAACTCCATCTTTCACGTAATGGCGAACCTGAAGGGAATTTTCAACAGA	99
QY	201	ACCTCAATCGAAGACTATAAGAAATGGCTTGTATTTCCATTTAGTGCTTCTTTGTCTCT	260
DB	100	AGAGAGAATCAAGAGTACTCTTTGGAGGTTAAGAGTGTCTCTATGTCACTCTCCTCTT	159
QY	261	TGCTTGCCAAAGCACTTTCTACAATTTTGGGCGAGTTTACTATGAAATGGTGGGAAGAG	320
DB	160	AGCTGAGAGACAATAGCCACTCTCTTAGTAGACTTTACTACGAAAAGCGGTAAAG	219
QY	321	TACATGGATGGGAACACTTGTGCAACTAATCGGCTTCCCTGTCTGTCTTCTTCCGCTT	380
DB	220	CACATGGCTCGAAACCTCTGGTTTCTAGGTTTCTTTAAACCTTCTCTTGTCTATTA	279
QY	381	CTTTTCCCAACCAAAATCCCAACCAACAGACAGATTTTCAGAAAGTTCTCTTCTCTT	440
DB	280	TTACTTAAAGCTTGAGCCGTCAAGACTTAAACCAATTCACAAAACACTACTTCTTCTT	339
QY	441	CACCAATCTTGGATCAGTTTACATCGTTTACTGGACTATTAGTGTCTGTCTTAACTCTTAT	500
DB	340	CTTGACACTATCTTTAGTGTATATGGACTTGGCTTGGCTTGTGTGGACATTTGATTTT	399
QY	501	GTCTCTGTGTGTTTACTATACTTACCAAGTTTCTACTTTCTCCCTCACTTGGCCTCACA	560
DB	400	GTACTCATTTGGGCTACTTTAGCTTCCCTCTCAACTTCTCTTTGTATCTCTGCGTCGCA	459
QY	561	ATTGGCTTCACTGGCTTCTCTCATATTTTCTAACTCGGAGAGTTTCACACCTTTCAT	620
DB	460	ATTGGCTTTTAAAGCGGCTTCTCTTACTTCTTAACTCACAATAACACACCACTTTAT	519
QY	621	TGTGAATTTCTGTGTTCTCTTACTATTCTCTGCTTCTGCTTCTGCTTCAACACTGATTC	680
DB	520	ACTCAATTCACCTTGTCTTCTTAAACATATCTTCTACACTTCTTGTATTCACATGAACC	579
QY	681	GGAA-----AACACAGCAAAAGTGTCTAGAGTAAATATGTGATAGGATAAT	728
DB	580	AGAAATCTCCCTCTTCTTACTTCAAAAGTCCGAGCAAGTCCAAAGTATGTGATTGGATACAT	639
QY	729	ATGTACCATTTGGTGTCTTCTGCTGGGATTTGGATTTGCTGTATCCCTGGTACAACTGATCT	788
DB	640	CTGCGCGGTCGGTAGCTAGCTAGCTGGTTATTTCTCTGCTGCTTCTTTTAAACAGATTACGCGTT	699
QY	789	CAGGAAGGTTTTTAAAGAACAACAATTTCTCAACGGTCACTGACTTGGTGGCTTTACCAATC	848
DB	700	CGAAAAGATCTTAAGAAATACACATTTCAAGGCTATTTTAGACATGGCCACATATCCGTC	759
QY	849	TCTAGTTGCAAGCTGTGTGGTTTCTCATAGGACTTTTTCGCAAGCGGGAGTGGAAAACCTTT	908

DB	760	TATGTTAGTACTTCTGTAGTTGTGTAGGACTTTTGTGAGTGGTGGTGGAAAAAGCT	819
QY	909	AACRAAGTGAGATGGAAAACTACAACTGGGAAAGTGCCATACGTTATGACTTTTGGCCTC	968
DB	820	GAGTACAGAAATGGAAGAGTTTCAACTAGGAAAGCTCATACATTTTGTATAAATCGG	879
QY	969	GATAGCTATTTCTCTGGCAAGTCTACACCAATTTGGCGTCTGGGACTGATCTTTGAGTCATC	1028
DB	880	TTCAACGATATCATGGCAAGCTTTTGTATTGGAAGTGTGTGTTGATTATCGAAGTTTC	939
QY	1029	TTCTGTGTTCTCCAAATTCATAAATGCTGCTGTGGGATTTGCTATAGTTCCAGTTGTAGCAGT	1088
DB	940	ATCGCTTTTCTCCAAATGTCATAAGCACTCTTTGTTTACAGTTGTGCTCTTTCTTCTGT	999
QY	1089	GATTCTTTTCCATGATAAAATGAACGCGTCAAGATCTTCTCCATCATTTTGTAGCTATCTG	1148
DB	1000	TGTCTTCTCCGTATGAGATGAGTGGAAATCAAGTTGATTGCAATGTTTGTGGCCATCTG	1059
QY	1149	GGGATTCATTTCTATGCTATCAGCACTACCTCGACGAAAGAAAG	1194
DB	1060	GGGATTTCTTCTTATGTTATCAGCAATATGTCATGATAGAAAG	1105
RESULT 12			
AAC47974			
ID	AAC47974 standard; DNA; 2660 BP.		
XX	AAC47974;		
XX	18-OCT-2000 (first entry)		
DE	Arabidopsis thaliana DNA fragment SEQ ID NO: 55799.		
KW	Hybridisation assay; genetic mapping; gene expression control;		
KW	protein identification; signal transduction pathway; metabolic pathway;		
XX	promoter; termination sequence; ss.		
OS	Arabidopsis thaliana.		
PN	EP1033405-A2.		
PD	06-SEP-2000.		
PF	25-FEB-2000; 2000EP-00301439.		
PR	25-FEB-1999; 99US-0121825P.		
PR	05-MAR-1999; 99US-0123180P.		
PR	09-MAR-1999; 99US-0123548P.		
PR	23-MAR-1999; 99US-0125788P.		
PR	25-MAR-1999; 99US-0126264P.		
PR	29-MAR-1999; 99US-0126785P.		
PR	01-APR-1999; 99US-0127462P.		
PR	06-APR-1999; 99US-0128234P.		
PR	16-APR-1999; 99US-0128714P.		
PR	19-APR-1999; 99US-0129845P.		
PR	21-APR-1999; 99US-0130077P.		
PR	23-APR-1999; 99US-0130510P.		
PR	23-APR-1999; 99US-0130891P.		
PR	28-APR-1999; 99US-0131449P.		
PR	30-APR-1999; 99US-0132048P.		
PR	30-APR-1999; 99US-0132407P.		
PR	04-MAY-1999; 99US-0132484P.		
PR	05-MAY-1999; 99US-0132485P.		
PR	06-MAY-1999; 99US-0132486P.		
PR	06-MAY-1999; 99US-0132487P.		
PR	07-MAY-1999; 99US-0132863P.		
PR	11-MAY-1999; 99US-0134256P.		
PR	14-MAY-1999; 99US-0134218P.		
PR	14-MAY-1999; 99US-0134219P.		
PR	14-MAY-1999; 99US-0134221P.		
PR	14-MAY-1999; 99US-0134370P.		

PR	18-MAY-1999;	99US-0134768P.
PR	19-MAY-1999;	99US-0134941P.
PR	20-MAY-1999;	99US-0135124P.
PR	21-MAY-1999;	99US-0135353P.
PR	24-MAY-1999;	99US-0135629P.
PR	25-MAY-1999;	99US-0136021P.
PR	27-MAY-1999;	99US-0136392P.
PR	28-MAY-1999;	99US-0136782P.
PR	01-JUN-1999;	99US-0137222P.
PR	03-JUN-1999;	99US-0137528P.
PR	04-JUN-1999;	99US-0137502P.
PR	07-JUN-1999;	99US-0137724P.
PR	08-JUN-1999;	99US-0138094P.
PR	10-JUN-1999;	99US-0138540P.
PR	12-JUN-1999;	99US-0138847P.
PR	14-JUN-1999;	99US-0139111P.
PR	16-JUN-1999;	99US-0139452P.
PR	16-JUN-1999;	99US-0139453P.
PR	17-JUN-1999;	99US-0139492P.
PR	18-JUN-1999;	99US-0139454P.
PR	18-JUN-1999;	99US-0139455P.
PR	18-JUN-1999;	99US-0139456P.
PR	18-JUN-1999;	99US-0139457P.
PR	18-JUN-1999;	99US-0139458P.
PR	18-JUN-1999;	99US-0139459P.
PR	18-JUN-1999;	99US-0139460P.
PR	18-JUN-1999;	99US-0139461P.
PR	18-JUN-1999;	99US-0139462P.
PR	18-JUN-1999;	99US-0139463P.
PR	18-JUN-1999;	99US-0139750P.
PR	18-JUN-1999;	99US-0139763P.
PR	21-JUN-1999;	99US-0139817P.
PR	22-JUN-1999;	99US-0139899P.
PR	23-JUN-1999;	99US-0140335P.
PR	23-JUN-1999;	99US-0140354P.
PR	24-JUN-1999;	99US-0140695P.
PR	28-JUN-1999;	99US-0140823P.
PR	29-JUN-1999;	99US-0140931P.
PR	30-JUN-1999;	99US-0141287P.
PR	01-JUL-1999;	99US-0141842P.
PR	01-JUL-1999;	99US-0142154P.
PR	02-JUL-1999;	99US-0142055P.
PR	06-JUL-1999;	99US-0142390P.
PR	08-JUL-1999;	99US-0142803P.
PR	09-JUL-1999;	99US-0142920P.
PR	12-JUL-1999;	99US-0142977P.
PR	13-JUL-1999;	99US-0143542P.
PR	14-JUL-1999;	99US-0143624P.
PR	15-JUL-1999;	99US-0144005P.
PR	16-JUL-1999;	99US-0144085P.
PR	16-JUL-1999;	99US-0144086P.
PR	19-JUL-1999;	99US-0144325P.
PR	19-JUL-1999;	99US-0144331P.
PR	19-JUL-1999;	99US-0144332P.
PR	19-JUL-1999;	99US-0144333P.
PR	19-JUL-1999;	99US-0144334P.
PR	19-JUL-1999;	99US-0144335P.
PR	20-JUL-1999;	99US-0144352P.
PR	20-JUL-1999;	99US-0144632P.
PR	20-JUL-1999;	99US-0144884P.
PR	21-JUL-1999;	99US-0144814P.
PR	21-JUL-1999;	99US-0145086P.
PR	21-JUL-1999;	99US-0145088P.
PR	22-JUL-1999;	99US-0145085P.
PR	22-JUL-1999;	99US-0145087P.
PR	22-JUL-1999;	99US-0145089P.
PR	22-JUL-1999;	99US-0145192P.
PR	23-JUL-1999;	99US-0145145P.
PR	23-JUL-1999;	99US-0145218P.
PR	23-JUL-1999;	99US-0145224P.
PR	26-JUL-1999;	99US-0145276P.
PR	27-JUL-1999;	99US-0145913P.
PR	27-JUL-1999;	99US-0145918P.

---

PR	27-JUL-1999;	99US-0145919P.
PR	28-JUL-1999;	99US-0145951P.
PR	02-AUG-1999;	99US-0146386P.
PR	02-AUG-1999;	99US-0146388P.
PR	02-AUG-1999;	99US-0146389P.
PR	03-AUG-1999;	99US-0147038P.
PR	04-AUG-1999;	99US-0147204P.
PR	04-AUG-1999;	99US-0147302P.
PR	05-AUG-1999;	99US-0147392P.
PR	06-AUG-1999;	99US-0147303P.
PR	06-AUG-1999;	99US-0147416P.
PR	09-AUG-1999;	99US-0147493P.
PR	09-AUG-1999;	99US-0147935P.
PR	10-AUG-1999;	99US-0148171P.
PR	11-AUG-1999;	99US-0148319P.
PR	12-AUG-1999;	99US-0148341P.
PR	13-AUG-1999;	99US-0148565P.
PR	13-AUG-1999;	99US-0148684P.
PR	16-AUG-1999;	99US-0149368P.
PR	17-AUG-1999;	99US-0149175P.
PR	18-AUG-1999;	99US-0149426P.
PR	20-AUG-1999;	99US-0149722P.
PR	20-AUG-1999;	99US-0149723P.
PR	20-AUG-1999;	99US-0149929P.
PR	23-AUG-1999;	99US-0149902P.
PR	23-AUG-1999;	99US-0149930P.
PR	25-AUG-1999;	99US-0150566P.
PR	26-AUG-1999;	99US-0150884P.
PR	27-AUG-1999;	99US-0151065P.
PR	27-AUG-1999;	99US-0151066P.
PR	27-AUG-1999;	99US-0151080P.
PR	30-AUG-1999;	99US-0151303P.
PR	31-AUG-1999;	99US-0151438P.
PR	01-SEP-1999;	99US-0151930P.
PR	07-SEP-1999;	99US-0152363P.
PR	10-SEP-1999;	99US-0153070P.
PR	13-SEP-1999;	99US-0153758P.
PR	13-SEP-1999;	99US-0154018P.
PR	16-SEP-1999;	99US-0154039P.
PR	20-SEP-1999;	99US-0154779P.
PR	22-SEP-1999;	99US-0155139P.
PR	23-SEP-1999;	99US-0155486P.
PR	24-SEP-1999;	99US-0155659P.
PR	28-SEP-1999;	99US-0156458P.
PR	29-SEP-1999;	99US-0156596P.
PR	04-OCT-1999;	99US-0157117P.
PR	05-OCT-1999;	99US-0157753P.
PR	06-OCT-1999;	99US-0157865P.
PR	07-OCT-1999;	99US-0158029P.
PR	08-OCT-1999;	99US-0158232P.
PR	12-OCT-1999;	99US-0158369P.
PR	13-OCT-1999;	99US-0159293P.
PR	13-OCT-1999;	99US-0159294P.
PR	13-OCT-1999;	99US-0159295P.
PR	14-OCT-1999;	99US-0159329P.
PR	14-OCT-1999;	99US-0159330P.
PR	14-OCT-1999;	99US-0159331P.
PR	14-OCT-1999;	99US-0159637P.
PR	14-OCT-1999;	99US-0159638P.
PR	18-OCT-1999;	99US-0159584P.
PR	21-OCT-1999;	99US-0160741P.
PR	21-OCT-1999;	99US-0160767P.
PR	21-OCT-1999;	99US-0160768P.
PR	21-OCT-1999;	99US-0160770P.
PR	21-OCT-1999;	99US-0160814P.
PR	21-OCT-1999;	99US-0160815P.
PR	22-OCT-1999;	99US-0160980P.
PR	22-OCT-1999;	99US-0160981P.
PR	22-OCT-1999;	99US-0160989P.
PR	25-OCT-1999;	99US-0161404P.
PR	25-OCT-1999;	99US-0161405P.
PR	25-OCT-1999;	99US-0161406P.







PR 21-JUN-1999;	99US-01339817P.	PR 27-AUG-1999;	99US-0151080P.
PR 22-JUN-1999;	99US-0139899P.	PR 30-AUG-1999;	99US-0151303P.
PR 23-JUN-1999;	99US-0140353P.	PR 31-AUG-1999;	99US-0151438P.
PR 23-JUN-1999;	99US-0140354P.	PR 01-SEP-1999;	99US-0151930P.
PR 24-JUN-1999;	99US-0140695P.	PR 07-SEP-1999;	99US-0152363P.
PR 28-JUN-1999;	99US-0140823P.	PR 10-SEP-1999;	99US-0153070P.
PR 29-JUN-1999;	99US-0140991P.	PR 13-SEP-1999;	99US-0153758P.
PR 30-JUN-1999;	99US-01411287P.	PR 15-SEP-1999;	99US-0154018P.
PR 01-JUL-1999;	99US-01411842P.	PR 16-SEP-1999;	99US-0154039P.
PR 02-JUL-1999;	99US-0142055P.	PR 20-SEP-1999;	99US-0154779P.
PR 06-JUL-1999;	99US-0142154P.	PR 22-SEP-1999;	99US-0155139P.
PR 08-JUL-1999;	99US-0142390P.	PR 23-SEP-1999;	99US-0155486P.
PR 09-JUL-1999;	99US-0142803P.	PR 24-SEP-1999;	99US-0155659P.
PR 12-JUL-1999;	99US-0142920P.	PR 28-SEP-1999;	99US-0156458P.
PR 13-JUL-1999;	99US-0143277P.	PR 29-SEP-1999;	99US-0156596P.
PR 14-JUL-1999;	99US-0143342P.	PR 04-OCT-1999;	99US-0157117P.
PR 15-JUL-1999;	99US-0143624P.	PR 05-OCT-1999;	99US-0157753P.
PR 16-JUL-1999;	99US-0144085P.	PR 06-OCT-1999;	99US-0157865P.
PR 16-JUL-1999;	99US-0144086P.	PR 07-OCT-1999;	99US-0158029P.
PR 19-JUL-1999;	99US-0144325P.	PR 08-OCT-1999;	99US-0158232P.
PR 19-JUL-1999;	99US-0144331P.	PR 12-OCT-1999;	99US-0158369P.
PR 19-JUL-1999;	99US-0144332P.	PR 13-OCT-1999;	99US-0159293P.
PR 19-JUL-1999;	99US-0144333P.	PR 13-OCT-1999;	99US-0159294P.
PR 19-JUL-1999;	99US-0144334P.	PR 13-OCT-1999;	99US-0159295P.
PR 19-JUL-1999;	99US-0144335P.	PR 14-OCT-1999;	99US-0159330P.
PR 20-JUL-1999;	99US-0144632P.	PR 14-OCT-1999;	99US-0159331P.
PR 20-JUL-1999;	99US-0144884P.	PR 14-OCT-1999;	99US-0159637P.
PR 21-JUL-1999;	99US-0144814P.	PR 14-OCT-1999;	99US-0159638P.
PR 21-JUL-1999;	99US-0145086P.	PR 18-OCT-1999;	99US-0159584P.
PR 21-JUL-1999;	99US-0145088P.	PR 21-OCT-1999;	99US-0160741P.
PR 22-JUL-1999;	99US-0145085P.	PR 21-OCT-1999;	99US-0160767P.
PR 22-JUL-1999;	99US-0145087P.	PR 21-OCT-1999;	99US-0160768P.
PR 22-JUL-1999;	99US-0145089P.	PR 21-OCT-1999;	99US-0160770P.
PR 22-JUL-1999;	99US-0145192P.	PR 21-OCT-1999;	99US-0160814P.
PR 23-JUL-1999;	99US-0145145P.	PR 21-OCT-1999;	99US-0160815P.
PR 23-JUL-1999;	99US-0145218P.	PR 22-OCT-1999;	99US-0160980P.
PR 23-JUL-1999;	99US-0145224P.	PR 22-OCT-1999;	99US-0160981P.
PR 26-JUL-1999;	99US-0145276P.	PR 22-OCT-1999;	99US-0160989P.
PR 27-JUL-1999;	99US-0145913P.	PR 25-OCT-1999;	99US-0161404P.
PR 27-JUL-1999;	99US-0145918P.	PR 25-OCT-1999;	99US-0161405P.
PR 28-JUL-1999;	99US-0145919P.	PR 25-OCT-1999;	99US-0161406P.
PR 28-JUL-1999;	99US-0145951P.	PR 26-OCT-1999;	99US-0161359P.
PR 02-AUG-1999;	99US-0146386P.	PR 26-OCT-1999;	99US-0161360P.
PR 02-AUG-1999;	99US-0146388P.	PR 26-OCT-1999;	99US-0161361P.
PR 03-AUG-1999;	99US-0146389P.	PR 28-OCT-1999;	99US-0161920P.
PR 03-AUG-1999;	99US-0147038P.	PR 28-OCT-1999;	99US-0161992P.
PR 04-AUG-1999;	99US-0147204P.	PR 28-OCT-1999;	99US-0161993P.
PR 04-AUG-1999;	99US-0147302P.	PR 29-OCT-1999;	99US-0162142P.
PR 05-AUG-1999;	99US-0147192P.		
PR 05-AUG-1999;	99US-0147260P.		
PR 06-AUG-1999;	99US-0147303P.		
PR 06-AUG-1999;	99US-0147416P.		
PR 09-AUG-1999;	99US-0147493P.		
PR 09-AUG-1999;	99US-0147935P.		
PR 10-AUG-1999;	99US-0148171P.		
PR 11-AUG-1999;	99US-0148319P.		
PR 12-AUG-1999;	99US-0148341P.		
PR 13-AUG-1999;	99US-0148565P.		
PR 13-AUG-1999;	99US-0148684P.		
PR 16-AUG-1999;	99US-0149368P.		
PR 17-AUG-1999;	99US-0149175P.		
PR 18-AUG-1999;	99US-0149426P.		
PR 20-AUG-1999;	99US-0149722P.		
PR 20-AUG-1999;	99US-0149723P.		
PR 23-AUG-1999;	99US-0149929P.		
PR 23-AUG-1999;	99US-0149902P.		
PR 23-AUG-1999;	99US-0149930P.		
PR 25-AUG-1999;	99US-0150566P.		
PR 26-AUG-1999;	99US-0150884P.		
PR 27-AUG-1999;	99US-0151065P.		
PR 27-AUG-1999;	99US-0151066P.		

Query Match

Best Local Similarity 17.8%; Score 230.8; DB 3; Length 1403;

Matches 518; Conservative 0; Mismatches 432; Indels 18; Gaps 1;

QY 226	TGGCTTCGTATTCCATTACGTGTTCTTTGTCCTTGTGTCCTTGGCCAAAGCAGCTTTTCTACAATT	285
Db		
275	TGGATTCTTGCTTTTATAAGTATCTTCTTCTCATCTCTGCTCAAGCCATTGCTGTCTT	334
QY 286	TTGGCAGAGTTTACTATGAAATGGTGGGAAGAGTACATGATGGGAACACTTGTCCAA	345
Db		
335	CTTGGTCGGTTTTTATTACACGAAGGTGGAACAGTAAATGGATCTCTACTCTTGTCCAA	394
QY 346	CTAATCGGCTCCCGTTTCTGTTTCTTTCGCTTCTTTTCCCAACCAAAAATCCCAA	405
Db		
395	ACTTGTGTTTTCCGATTCTCTATCTCCCTCTTTGTTTCTTCTTCTTCACTCTTCT	454
QY 406	CCACAGAAAGCAGATTTTCAGAAAGTTCTTCTTCCCTTCCACATTTCTTGGATCAGTTTACATC	465
Db		
455	TCATCT-----TCCTGTTCTTCAAGACTCTGGTTTGGATTATCTT	496
QY 466	GTACTGGAATTTAGTGTCTGCTAACTCTTATATGTCCTCTGTGGTTTACTATACTTA	525
Db		
497	TCTCTTGGTTTTGCTATTGGTTTAGACAAATCTTTTATACTCTTTTGGTCTTTTGTATCTC	556

```
Qy 526 CCAGTTTCTACTTCTCCCTCATCTTGGCCCTCACAATTGGCCCTTCACTGCCTTCTTCTCA 585
Db 557 TCTGCTTCAACTATTCGATCTTGTGTTCTTTCACAGTTAGCTTTCAATGGTGTCTTCTCT 616
Qy 586 TATTCTTAAACTCGCAGAAGTTCACACCTTTTCATTGTGAATTTCTGTTTCTCTTCTTACT 645
Db 617 TATTACATCAATTCAGAAAATCATTGTTGATCTCTTCTCTGTGTTGTTTCTCTCT 676
Qy 646 ATTTCCTCTGCCCTCTCGTGGTCAACACTGATTTCGGAAAAACACAGCAAAAGTCTCAGA 705
Db 677 GTCCTGCTGTGTTGTTCTCTTGACGATGATTCCAATAGCCATCAGGAGATTCCTAAG 736
Qy 706 GTAAATATGATAGGATATATATGATACATTGGTCTTCTGCTGGATGGATTGCTG 765
Db 737 TGGAGTACTTATGATGTTGTTTATGATCTGTTTGTCTTCTTATCTATCTTCTTCAG 796
Qy 766 CTATCCCTGGTACAACTGATCCTCAGGAAGTTTTTAAAGAACCAACATTCTCAACGTC 825
Db 797 CTCCTCTTATGCACTTTCTTTCAGAAATGTTCTCAAGATGAGACGTTCTCTATGGTT 856
Qy 826 ACTGACTTGGTGGCTTACCAATCTCTAGTTGCAAGCTGTGTGGTTCTCATAGGACTTTTC 885
Db 857 CTCGAGATGCAAACTATACGCTGCTTGTGGCTTCTTGTGTTGCGGTTATCGGTTGTT 916
Qy 886 GCAAGCGGGAGTGGAAAACCTTTAACAAGTCAGATGGAAACTCAAACTGGGGAAGTG 945
Db 917 GCGAGCGGGGAATGGATGTTGTGAGTGGAGATGGGAAGATTTTCACGAGGGTCAAGTC 976
Qy 946 CCATACGTTATGACTTTGGCCTCGATAGCTATTTCTCGCAAGTCTACACCATTTGGCGTC 1005
Db 977 ATTTATGTTTGAATTTGTTGGACCGCGGTTTGTGGCAATTTGGGTTCTGTAGAGCC 1036
Qy 1006 GTGGACTGATCTTTGAGTCATCTCTGTGTTCTTCCAAATCCATAACTGCTGTGGGATG 1065
Db 1037 GTGGCGCTTATATTTCTGTTCTTCTGCTGTTTTCGAACCTTATTTGGTACACTCTCACT 1096
Qy 1066 CCTATAGTTCAGTTGTAGCAGTATGTTTTCATGATATAAATGAACGGCTCAAGATC 1125
Db 1097 ATTTGTACGCTCTTTGCAGCCATTTGCGGTGTTCCACGACAAAGCTGAGGTTTAAAGATG 1156
Qy 1126 TTCTCCATCATTTAGCTATCTGGGATTCATTTTCTATGCTATCAGCACTACTCTGAC 1185
Db 1157 GTCGGATGCTCATTTGCCCTTATGGGGTTTCGGGTTTATATCTACCAGAAATTTCTTGAT 1216
Qy 1186 GAAAGAA 1193
Db 1217 GACCTGAA 1224
RESULT 15
ADA70748
ID ADA70748 standard; DNA; 2175 BP.
XX AC ADA70748;
XX 20-NOV-2003 (first entry)
XX DE Rice gene, SEQ ID 4071.
XX KW Plant; bacterial infection; fungal infection; viral infection; rice;
XX OS gene; db.
XX PN Oryza sativa.
XX PD WO200300898-A1.
XX PF 03-JAN-2003.
XX PR 22-JUN-2001; 2001WO-IB001105.
XX 22-JUN-2001; 2001WO-IB001105.
```

(SVGN ) SYNGENTA PARTICIPATIONS AG.

Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;  
Katagiri F, Qian S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;  
WPI; 2003-175290/17.

Identifying at least one gene involved in plant resistance or response to  
pathogenic infection for conferring resistance or tolerance to a plant to  
bacterial, fungal or viral infection by determining or detecting plant  
gene expression.

Claim 6; SEQ ID NO 4071; 899pp; English.

The present invention relates to a method (M1) for identifying genes  
involved in plant resistance or response to pathogenic infection. M1  
comprises identifying a gene whose expression is significantly altered in  
the incompatible interaction of plant gene expression relative to  
expression of the gene in an uninfected plant, in a mutant plant that  
does not express a gene associated with response to pathogenic infection,  
or in a corresponding incompatible or compatible interaction. (M1) is  
useful for conferring resistance to resistance or tolerance to a plant to  
bacterial, fungal or viral infection. The present sequence was used to  
illustrate the invention.

Sequence 2175 BP; 363 A; 701 C; 671 G; 439 T; 0 U; 1 Other;

Query Match 17.6%; Score 227.2; DB 8; Length 2175;  
Best Local Similarity 54.9%; Pred. No. 1e-56;  
Matches 479; Conservative 0; Mismatches 378; Indels 15; Gaps 1;

```
Qy 245 ACCTGTTCTTGTCTTGTGCTTGCACAGCACTTTCTACAAATTTGGGAGATTACTATG 304
Db 158 ACCTGTTCTTCTTCTATCGCGGTACAGATCGGCGACACTGCTGGCGAGATACTACTCA 217
Qy 305 AAAATGTGTGGAAGATACATGGATGGGAACACTTGTCCAACTAAATCGGCTTCCCTGTC 364
Db 218 CCNAGCGCGCCAGCAAGTGGATATCGGGTTCGTGCGGACGGCGGCTCCCGATAC 277
Qy 365 TGTTCCTTTCGCTTCTTTTCCAAACCAAAATCCCAAAACCAACAGAGAGATTCA 424
Db 278 TGTTCCTTTCACCTTCTTCTTCTTCCCTCGAAGTCACCGCTCTC-----CT 322
Qy 425 GAAAGTCTCTCTTCCACCATTTCTGGATCAGTTTACATCGTTACTGGACTATTAGTGT 484
Db 323 GCACAAACACTCCCATGCGCAAGCTCCCGTGATATACATCGTCTGGGCTCATCATCG 382
Qy 485 CTGCTAACTCTTATATGCTCTTGTGTTTACTATATCTTACCAGTTTCTACTTTCTCCC 544
Db 383 CGCGCGACGACATGATGATCACTGGTGGCTCAAGTACCTCCCGCTTCGACCTACTCGC 442
Qy 545 TCATCTTGGCTCACAATGGCCTTCACTGCTTCTTCTCTCATATTTTCTAACTCGGAGA 604
Db 443 TCATCTGTCAGCCAGCTCGCGTTCAATGTGCTTCTCTCATAGCTGCTCAACTCCCGA 502
Qy 605 AGTTACACCTTTTATTGTTGAATTTCTGTTTCTTCTTCTTCTTCTTCTGCTCTCTCG 664
Db 503 AGGTCACTCTCTGATATTCAACTCGCTGCTGCTGCTCACAATGCTCGCTTCGCTCATCG 562
Qy 665 TGGTCAACACTGATTCGGAACAAACACAGCAAAAGTGTCTAGAGTAAATATGTGATAGGA 724
Db 563 GAGTCAGCAAGGAGTCTCAGGGGGTCAACCGCGTCTCGGGAGGGAAGTATCTGCTGTT 622
Qy 725 TAATATGATACCATTTGGTGTCTTCTGCTGGGATGGAATGCTGCTATCCCTGGTCAACTGA 784
Db 623 TCGTGTGACGCTGGGGCGTCTGTCACCTACTCGCTGATCTCTGCGCTCATGAGCTCA 682
Qy 785 TCCTCAGGAAGGTTTTAAAGAACCAACATTCTCAACGGTCACGTGCTGCTCTTACC 844
Db 683 CTTTCAGAGCCATCATCAAGAAGCACACCTTCTCAGCCGCTCTCAACATGAGATCTACA 742
Qy 845 AATCTCTAGTTGCAAGCTGTGTGTTCTCTCATAGACTTTTTCGCAAGCGGGAGTGGAAA 904
```

Db	743	CGGGCTCGTGGCGACGGCCGCGTCGGTGTGGGCTGTTCGCGAGCGGCGAGTGGAGGT	802
Qy	905	CTTTAACAGTAGAGATGGAATACTACAAACTGGGGAAAGTGCCATACGTTATGACTTTGG	964
Db	803	CGCTGAGGGGGAGATGAACGCGTTCAGGTCGGGGCAGTTCTCTACCTGATGACGCTGC	862
Qy	965	CCTCGATAGCTATTTCTCTGGAAGTCTACACCAATGGCGTCGTGGGACTGATCTTTGAGT	1024
Db	863	TGTGGGCGCGCTGTCTGTGGCAGGTGGCCAAACATCGGGGTGCTCGGCCCTCATCTTCGAGG	922
Qy	1025	CATCTTCTGTGTCTCCAAATTCATAACTGCTGTGGGATGGCCTATAGTCCAGTTGTAG	1084
Db	923	TGTGGGCGCTCTTCTCCAAAGTATCAGCAGGTGTCACATGCGGTCATCCCGTCTCTCG	982
Qy	1085	CAGTGATTTGTTTCCATGATAAAATGAACGCG	1116
Db	983	CGGTGGTGGTGTTCACGACAGGATGAACGGG	1014

Search completed: November 1, 2004, 07:19:55  
Job time : 684.972 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 1, 2004, 02:44:43 ; Search time 5730.74 Seconds  
(without alignments)  
10669.751 Million cell updates/sec

Title: US-09-913-767-4  
Perfect score: 1293  
Sequence: 1 ggaagtcctcttagtggtggt.....tatgatcaaaacatatctcc 1293

Scoring table: IDENTITY NUC  
Gapop 10\_0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues  
Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*  
1: gb\_ba.\*  
2: gb\_htg.\*  
3: gb\_in.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_ete.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match Length DB ID	Description
1	1293	100.0 1293 6 AX033547	AX033547 Sequence
2	1293	100.0 58427 8 ATF15J5	AL110123 Arabidops
3	1293	100.0 82697 8 ATT9A21	AL021713 Arabidops
4	1293	100.0 194143 8 ATCHRIV48	AL161548 Arabidops
5	1152	89.1 3387 6 AX507517	AX507517 Sequence
6	1152	89.1 3387 6 AX652015	AX652015 Sequence
7	816.6	63.2 1194 6 AX033548	AX033548 Sequence
8	775	59.9 1165 8 AY096558	AY096558 Arabidops
9	775	59.9 1428 8 AY074546	AY074546 Arabidops
10	454.4	35.1 1106 8 AF370622	AF370622 Arabidops
11	451.8	34.9 733 8 AF083706	AF083706 Arabidops
12	435.2	33.7 1071 6 AX033550	AX033550 Sequence
13	429.6	33.2 721 11 BX664674	BX664674 Arabidops
14	410	31.7 1047 6 AX412868	AX412868 Sequence
15	410	31.7 1081 6 AX033549	AX033549 Sequence
16	390.8	30.2 73179 8 AP006383	AP006383 Lotus cor
17	295	22.8 1617 8 AK072660	AK072660 Oryza sat
18	295	22.8 175645 8 AP004096	AP004096 Oryza sat
19	294.4	22.8 1145 6 AX033546	AX033546 Sequence

20	293.4	22.7	1476	8	AK071254	AK071254 Oryza sat
21	289	22.4	1251	8	BT003358	BT003358 Arabidops
22	289	22.4	1372	8	AY062782	AY062782 Arabidops
23	289	22.4	115421	8	T12C22	AC020576 Sequence
24	282.4	21.8	110036	8	AC142095	AC142095 Medicago
25	269	20.8	133299	8	OSUN00098	AL606658 Oryza sat
26	269	20.8	145739	8	OSUN00157	AL662957 Oryza sat
27	265.8	20.6	1486	8	AK119483	AK119483 Oryza sat
28	246.6	19.1	83698	6	AX540653	AX540653 Sequence
29	246.6	19.1	83698	6	AB010072	AB010072 Arabidops
30	245.2	19.0	1279	8	BT014359	BT014359 Lycopersi
c 31	234	18.1	104386	8	AT132A17	AL161813 Arabidops
c 32	234	18.1	199362	8	ATCHRIV24	AL161512 Arabidops
33	230.8	17.8	1117	8	AY096468	AY096468 Arabidops
34	230.8	17.8	1333	8	AY074535	AY074535 Arabidops
35	230.8	17.8	1403	8	AY087844	AY087844 Arabidops
36	227.2	17.6	2175	6	AX654201	AX654201 Sequence
37	214.4	16.6	1152	6	AX653666	AX653666 Sequence
38	179	13.8	1361	8	AK071680	AK071680 Oryza sat
39	157	12.1	1049	6	AX033545	AX033545 Sequence
40	140.6	10.9	1155	6	AX653139	AX653139 Sequence
41	140.6	10.9	1599	8	AK060241	AK060241 Oryza sat
42	128.4	9.9	1102	8	BT005504	BT005504 Arabidops
43	128.4	9.9	1250	8	AK117664	AK117664 Arabidops
44	126.8	9.8	1225	6	AX033544	AX033544 Sequence
45	126.8	9.8	1227	8	AF078531	AF078531 Arabidops

ALIGNMENTS

RESULT 1  
AX033547  
LOCUS AX033547 1293 bp DNA linear PAT 21-SEP-2000  
DEFINITION Sequence 4 from Patent WO049152.  
ACCESSION AX033547  
VERSION AX033547.1 GI:10280291  
KEYWORDS  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana

Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsais.  
1  
Andre B., Buerkle, L., Frommer, W.B. and Gillissen, B.  
Nucleic acids that code for a nucleobase transporter  
Patent: WO 049152-A 4 24-AUG-2000;  
ANDRE BRUNO (BE) ; BUERKLE LUKAS (DE) ; FROMMER WOLF B (DE) ;  
GILLISSEN BERND (DE)

FEATURES  
source  
1..1293  
/organism="Arabidopsis thaliana"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:3702"

ORIGIN			
Query Match	100.0%	Score 1293;	DB 6; Length 1293;
Best Local Similarity	100.0%;	Pred. No. 0;	
Matches 1293;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;
Qy	1	GGAAGTCCTCTTAGAGTGGTTCATGGAATAACTCAAGTAATCTATGTCAATGGTAAGCAA	60
Db	1	GGAAGTCCTCTTAGAGTGGTTCATGGAATAACTCAAGTAATCTATGTCAATGGTAAGCAA	60
Qy	61	GATGCATCTCGTAGTGTAGATTACTTGTTCGCTAACCTGTTGTTTGTATGTTT	120
Db	61	GATGCATCTCGTAGTGTAGATTACTTGTTCGCTAACCTGTTGTTTGTATGTTT	120
Qy	121	TCAGATCATAACAATAGAACCACTACAGTCTAGGAGGAATAAATACCAATCCACCATG	180
Db	121	TCAGATCATAACAATAGAACCACTACAGTCTAGGAGGAATAAATACCAATCCACCATG	180
Qy	181	GAATTCGAATCTTCGTCGCTACCTCAATCGAAGAACTATAAGAAATGCTTCGTATTTCC	240

Db	181	GAATCGAATCTTCTGTCGTAACCTCAATCGAAGAACTATAGAATGCGTTCGATTTC	240
Qy	241	ATTTACGTGTTCTTTGTCCTTGTGCAAGCACTTTCTACAATTTTGGGAGAGTTTAC	300
Db	241	ATTTACGTGTTCTTTGTCCTTGTGCAAGCACTTTCTACAATTTTGGGAGAGTTTAC	300
Qy	301	TATGAAATGTTGGGAAGATACATGGATGGGAACACTTGTCCAACTAATCGGCTTCCT	360
Db	301	TATGAAATGTTGGGAAGATACATGGATGGGAACACTTGTCCAACTAATCGGCTTCCT	360
Qy	361	GTCTGTTCTCTTCCGCTCTCTTTCCCAACCAAAATCCCAACCAAGAGAGAT	420
Db	361	GTCTGTTCTCTTCCGCTCTCTTTCCCAACCAAAATCCCAACCAAGAGAGAT	420
Qy	421	TTGAAAGTTCTCTTCCCTCCACCATCTTGGATCAGTTTACATCGTTACTGGACTATTA	480
Db	421	TTGAAAGTTCTCTTCCCTCCACCATCTTGGATCAGTTTACATCGTTACTGGACTATTA	480
Qy	481	GTGCTGTAACCTCTTATATGTCCTCTGTTGGTTTACTATACCAAGTTTCTACTTTC	540
Db	481	GTGCTGTAACCTCTTATATGTCCTCTGTTGGTTTACTATACCAAGTTTCTACTTTC	540
Qy	541	TCCCTCATCTTGGCTCAAAATGGCCCTTCACTGCTCTTCTCATATTTTCTAAACTCG	600
Db	541	TCCCTCATCTTGGCTCAAAATGGCCCTTCACTGCTCTTCTCATATTTTCTAAACTCG	600
Qy	601	CAGAAGTTACACCTTTTCATTTGTAATTTCTCTGTTTCTCTTACTATTTCCTGCGCTC	660
Db	601	CAGAAGTTACACCTTTTCATTTGTAATTTCTCTGTTTCTCTTACTATTTCCTGCGCTC	660
Qy	661	CTCGTGTCACACTGATTCGAAAAACACAGCAAAAGTGTCTAGAGTAAATATGTGATA	720
Db	661	CTCGTGTCACACTGATTCGAAAAACACAGCAAAAGTGTCTAGAGTAAATATGTGATA	720
Qy	721	GGGATAATATGTACCATTTGTTCTCTGCTGGATTTGATGCTGCTATCCTGCTACAA	780
Db	721	GGGATAATATGTACCATTTGTTCTCTGCTGGATTTGATGCTGCTATCCTGCTACAA	780
Qy	781	CTGATCTCTCAGGAAGTTTAAAGAACAAACATTTCTCAACGGTCACTGACTTGTGCT	840
Db	781	CTGATCTCTCAGGAAGTTTAAAGAACAAACATTTCTCAACGGTCACTGACTTGTGCT	840
Qy	841	TACCAATCTTAGTTGCAAGTGTGTGTTCTCATAGGACTTTTCGCAAGCGGAGTGG	900
Db	841	TACCAATCTTAGTTGCAAGTGTGTGTTCTCATAGGACTTTTCGCAAGCGGAGTGG	900
Qy	901	AAAACCTTTAAACAAGTGAGATGAAACCTACAAACCTGGGAAAGTGCCATACGTTATGACT	960
Db	901	AAAACCTTTAAACAAGTGAGATGAAACCTACAAACCTGGGAAAGTGCCATACGTTATGACT	960
Qy	961	TTGGCTCGATAGCTATTTCCTGGCAAGTCTACACCAATGGCGTGGGGACTGATCTTT	1020
Db	961	TTGGCTCGATAGCTATTTCCTGGCAAGTCTACACCAATGGCGTGGGGACTGATCTTT	1020
Qy	1021	GAGTCATCTTCTGTTCTTCCAAATCCATACTGCTGTGGATTTGCTATAGTTCCAGTT	1080
Db	1021	GAGTCATCTTCTGTTCTTCCAAATCCATACTGCTGTGGATTTGCTATAGTTCCAGTT	1080
Qy	1081	GTAGCAGTATTGTTTTCATGATAAATGAACGGTCAAGAATCTTCTCCATCATTTTA	1140
Db	1081	GTAGCAGTATTGTTTTCATGATAAATGAACGGTCAAGAATCTTCTCCATCATTTTA	1140
Qy	1141	GCTATCTGGGATTCATTTTCTATCAGCACTACCTCGAGAAAGAGTTGAAG	1200
Db	1141	GCTATCTGGGATTCATTTTCTATCAGCACTACCTCGAGAAAGAGTTGAAG	1200
Qy	1201	ACTAGCCACAGTCTCTGATGAGATCTCTATCTACTACCTGCTGAGAAAGTTCACACA	1260
Db	1201	ACTAGCCACAGTCTCTGATGAGATCTCTATCTACTACCTGCTGAGAAAGTTCACACA	1260
Qy	1261	AACATACATAGTATGATCAAAACATATTTCC	1293

Db	1261	AACATACATAGTATGATCAAAACATATTTCC	1293
RESULT 2			
ATF15J5			
LOCUS	ATF15J5	58427 bp	DNA linear
DEFINITION	Arabidopsis thaliana DNA chromosome 4, BAC clone F15J5 (ESSA project).		PLN 30-AUG-1999
ACCESSION	AL110123		
VERSION	AL110123.1	GI:5816989	
KEYWORDS			
SOURCE	Arabidopsis thaliana (thale cress)		
ORGANISM	Arabidopsis thaliana		
REFERENCE			
AUTHORS	Bevan, M., Murphy, G., Ridley, P., Hudson, S., Bancroft, I., Mewes, H.W., Mayer, K.P.X., Lemcke, K. and Schueller, C.		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 58427)		
AUTHORS	EU Arabidopsis sequencing project.		
TITLE	Direct Submission		
JOURNAL	Submitted (30-MAR-1999) MIPS, at the Max-Planck-Institut fuer Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail: schuellemips.biochem.mpg.de, mayer@mips.biochem.mpg.de Project Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK, E-mail: michael.bevan@brc.ac.uk		
COMMENT	Information on performance of analysis and a more detailed annotation of this entry and other sequences of chromosomes 3, 4 and 5 can be viewed at: <a href="http://www.mips.biochem.mpg.de/proj/thal/">http://www.mips.biochem.mpg.de/proj/thal/</a> .		
FEATURES			
source	Location/Qualifiers		
	1. .58427		
	/organism="Arabidopsis thaliana"		
	/mol_type="genomic DNA"		
	/variety="Columbia"		
	/db_xref="taxon:3702"		
	/chromosome="4"		
misc_feature	1. .2548		
	/note="overlap to BAC T6K21; please refer to EMBL:AL021889 for analysis and annotation"		
gene	3596..5023		
	/gene="F15J5.10"		
gene	complement(join(3596..3646,3736..3801,3895..4020,4116..4281,4725..5023))		
	/gene="F15J5.10"		
CDS	complement(join(3596..3646,3736..3801,3895..4020,4116..4281,4725..5023))		
	/gene="F15J5.10"		
	/note="Contains Eukaryotic initiation factor 4E signature [DPYCFKHIEPKWEDPICANGKW] contains EST gb:T21435, AA586111"		
	/codon_start=1		
	/product="translation initiation factor eIF4E"		
	/protein_id="CA853645.1"		
	/db_xref="GI:5816990"		
	/db_xref="GOA:O23252"		
	/db_xref="Swiss-Prot:O23252"		
	/translation="MAVEDTPKSVTTEAKPNSIENPIDRYHEEGDDAEERGEIAGGEG DGNVDSKSGVPESHPLSHWTFFNDNPAVKSKOTSGSSLRPVTFTSVIEFWSLY NNKHKPSKLAHGAADYCFKHIIEPKWEDPICANGKWMTFPKEKSDKSWLYTLALIL GEQFDHGEDECAGVNIIRKQERISITWKNASNEAAQVSIQKQKWFELDYNNSIGFII HEDAKKLDRNAKNAYTA"		
exon	complement(3596..3646)		
	/gene="F15J5.10"		
	/number=1		
intron	complement(3647..3735)		
	/gene="F15J5.10"		
	/number=1		
exon	complement(3736..3801)		
	/gene="F15J5.10"		
	/number=2		





Db 56690 TCAGATCATAACAACATGAAGCAAAACCTAACAGCTCAGGAGGAATGAATACCAACATG 56749  
Qy 181 GAAATCGAATCTTCGTCCGTACCTCAATCGAAGAACTATAAGAAATGCTTCGTATTCC 240  
Db 56750 GAAATCGAATCTTCGTCCGTACCTCAATCGAAGAACTATAAGAAATGCTTCGTATTCC 56809  
Qy 241 ATTTACGTGTTCTTTGCTCCTTTCGCAAGCAGCTTTCTACAATTTGGGAGAGTTTAC 300  
Db 56810 ATTTACGTGTTCTTTGCTCCTTTCGCAAGCAGCTTTCTACAATTTGGGAGAGTTTAC 56869  
Qy 301 TATGAAATGTTGGGAGAGTACATGATGGATGGAACTTTGTCCAACTAATCGGCTTCCT 360  
Db 56870 TATGAAATGTTGGGAGAGTACATGATGGATGGAACTTTGTCCAACTAATCGGCTTCCT 56929  
Qy 361 GTTCTGTTCTCTTCGCGCTCTTTTCCAAACCAAAATCCCAACCAAGAGCAGAT 420  
Db 56930 GTTCTGTTCTCTTCGCGCTCTTTTCCAAACCAAAATCCCAACCAAGAGCAGAT 56989  
Qy 421 TTCAGAAAGTTCTCTTCCTTCACCAATCTTGATCAGTTTACATCGTTACTGGACTATTA 480  
Db 56990 TTCAGAAAGTTCTCTTCCTTCACCAATCTTGATCAGTTTACATCGTTACTGGACTATTA 57049  
Qy 481 GTGCTGCTAACTCTTATATGTCCTCTGTTGTTTACTATCTTACCAGTTTCTACTTTTC 540  
Db 57050 GTGCTGCTAACTCTTATATGTCCTCTGTTGTTTACTATCTTACCAGTTTCTACTTTTC 57109  
Qy 541 TCCCTCATCTTGGCTCACAATGGCCCTTCACTGCCCTTCTCTCATATTTTCTAAACTCG 600  
Db 57110 TCCCTCATCTTGGCTCACAATGGCCCTTCACTGCCCTTCTCTCATATTTTCTAAACTCG 57169  
Qy 601 CAGAAGTTACACCTTTCAATGTGAATCTCTGTTTCTCCTTACTATTTCTCTCGCCCTC 660  
Db 57170 CAGAAGTTACACCTTTCAATGTGAATCTCTGTTTCTCCTTACTATTTCTCTCGCCCTC 57229  
Qy 661 CTGCTGGTCAACACTGATTCGGAAACACACCAAAAGTCTCAGAGTAAATATGTGATA 720  
Db 57230 CTGCTGGTCAACACTGATTCGGAAACACACCAAAAGTCTCAGAGTAAATATGTGATA 57289  
Qy 721 GGGATAATATGATACATTGGTGCTTCTGCTGGGATTTGGATGCTATCCCTCGGTACAA 780  
Db 57290 GGGATAATATGATACATTGGTGCTTCTGCTGGATTTGGATGCTATCCCTCGGTACAA 57349  
Qy 781 CTGATCTCAGAGAGTTTAAAGAGCAAAATCTCAACGGTCACTGACTTGGTGGCT 840  
Db 57350 CTGATCTCAGAGAGTTTAAAGAGCAAAATCTCAACGGTCACTGACTTGGTGGCT 57409  
Qy 841 TACCAATCTAGTTGCAAGCTGTGTGTTCTCTAGGACTTTTCGCAAGCGGGAGTGG 900  
Db 57410 TACCAATCTAGTTGCAAGCTGTGTGTTCTCTAGGACTTTTCGCAAGCGGGAGTGG 57469  
Qy 901 AAAAATTTAACAAGTGAGATGAAACTACAACTGGGAAAGTGCCATAGTTTATGACT 960  
Db 57470 AAAAATTTAACAAGTGAGATGAAACTACAACTGGGAAAGTGCCATAGTTTATGACT 57529  
Qy 961 TTGGCTCGATAGCTATTTCTTGGCAAGTCTACACCATTTGGCGTGGGACTGATCTTT 1020  
Db 57530 TTGGCTCGATAGCTATTTCTTGGCAAGTCTACACCATTTGGCGTGGGACTGATCTTT 57589  
Qy 1021 GAGTCATCTTCTGTGTTCTCCAAATTCATAACTGCTGTGGGATTCCTATAGTTTCCAGTT 1080  
Db 57590 GAGTCATCTTCTGTGTTCTCCAAATTCATAACTGCTGTGGGATTCCTATAGTTTCCAGTT 57649  
Qy 1081 GTACAGTGATTTGTTTCCATGATAAAATGAACCGCTCAAGAGTCTTCTCCATCATTTTA 1140  
Db 57650 GTACAGTGATTTGTTTCCATGATAAAATGAACCGCTCAAGAGTCTTCTCCATCATTTTA 57709  
Qy 1141 CCTATCTGGGATTCATTTTCTATGTTGTTATCAGCACTTACCTCGACGAAAGAGTTGAAG 1200  
Db 57710 GCTATCTGGGATTCATTTTCTATGTTGTTATCAGCACTTACCTCGACGAAAGAGTTGAAG 57769  
Qy 1201 ACTAGCCACACAAGTCTGTAGGAGATCCTCATCTTACTACCTGTGAGGAGGTCACACA 1260

Db 57770 ACTAGCCACACAAGTCTCTGTAGGAGATCCTCATCTACTACTGCTGAGGAGGTCACACA 57829  
Qy 1261 AACATACATAGTGTATGATCAAAACATATTTC 1293  
Db 57830 AACATACATAGTGTATGATCAAAACATATTTC 57862  
RESULT 3  
ATT9A21  
LOCUS Arabidopsis thaliana DNA 82697 bp linear PLN 23-SEP-1999  
DEFINITION project).  
ACCESSION AL021713  
VERSION AL021713.1 GI:2832689  
KEYWORDS  
SOURCE  
ORGANISM Arabidopsis thaliana (thale cress)  
Arabidopsis thaliana  
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.  
REFERENCE 1  
AUTHORS Bevan, M., Murphy, G., Ridley, P., Hudson, S., Bancroft, I., Mewes, H.W.,  
Mayer, K.P.X., Lemcke, K. and Schueller, C.  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 82697)  
AUTHORS EU Arabidopsis sequencing, project.  
TITLE Direct Submission  
JOURNAL Submitted (22-SEP-1999) MIPS, at the Max-Planck-Institut fuer  
Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail:  
schueller@mps.biochem.mpg.de, mayer@mps.biochem.mpg.de Project  
Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge  
Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK,  
E-mail: michael.bevan@bbsrc.ac.uk  
COMMENT Information on performance of analysis and a more detailed  
annotation of this entry and other sequences of chromosomes 3, 4  
and 5 can be viewed at: <http://www.mips.biochem.mpg.de/proj/thal/>.  
FEATURES  
source  
1. 82697  
/organism="Arabidopsis thaliana"  
/mol\_type="genomic DNA"  
/variety="Columbia"  
/db\_xref="taxon:3702"  
/chromosome="4"  
402..1585  
/gene="T9A21.10"  
join(402..832,946..1090,1178..1585)  
/gene="T9A21.10"  
/notes="similarity to SPF1 protein, sweet protein,  
PIR2:S51529 and WRKY protein family, Petroselinum crispum,  
MNOS:S72443, MNOS:S72444, MNOS:S72445"  
/codon\_start=1  
/product="DNA binding-like protein"  
/protein\_id="CAA16788.1"  
/db\_xref="GI:4038481"  
/db\_xref="GOA:O49720"  
/db\_xref="GOA:Q8VWJ2"  
/db\_xref="Swiss-Prot:Q8VWJ2"  
/translation="MSNETRDLNYQYPSFSFLHEMNLFTSPSSYGNLPSQNGFNP  
STYFSDICLQSSPAAYELSLQFTGLSPSSSEVFNSSIDQEPNDRVTDVINGACNE  
TETRVSPNSSSSEADHPGDSKGRKRELVGEEQIISKVKGTKKTEVKQREPRV  
SPTKSEVDHLEDGVRWKYGOKAVKNSPYPRIIANGNENRSYYRCTOKCNVKKRVE  
RSQDPTVITTYTGGHNPITNLRGSSAAAFSADLMTSPSFAHDMPTAAYNG  
GSVAALADYVGSGSGVSNPSHSHQVTHQGEYELLREIFFSIFPKQEP"  
402..832  
/gene="T9A21.10"  
/number=1  
833..945  
/gene="T9A21.10"  
/number=1  
946..1090  
/gene="T9A21.10"  
/number=2  
1091..1177  
exon  
intron  
exon  
intron

exon /gene="T9A21.10"  
/number=2  
1178. .1585  
/gene="T9A21.10"  
/number=3  
4778. .6211  
/genes="T9A21.20"  
join(<4778. .4921,5023. .5187,5263. .5552,5638. .5866,  
5951. .6211)  
/genes="T9A21.20"  
/note="strong similarity to exopolysaccharuronase,  
Arabidopsis thaliana, PIR2:S34266"  
/codon\_start=1  
/product="polygalacturonase-like protein"  
/protein\_id="CAA16789.1"  
/db\_xref="GI:2832691"  
/db\_xref="GOA:O49721"  
/db\_xref="TrEMBL:O49721"  
/translation="AFVAWDKACKSSSSVNLIIIPRGFEVSGSRFSGPCNTVNSLT  
VRKATDLSKYRSGGWIQFGINGLTITGGTFDQGALAWFNNTCTSDSNCKLLP  
TSLKFGVMRTVVRRISSVNSKFHIALVECDKFGTRLNITAPSDSPNTDGIHRS  
SNVYSRSHIATGDDCVSIGQNSQITITSIKCGPHGHSVGLGRYPNEKOVNGLV  
KCKISGTTNGIRIKTWANSPGLSAATNMTFENIMNNVTNPILIDQSCPESSCISN  
VPSKVELSIYFKNIRGTSSSLVAVQLHCSRGMPCKKVVLENVHLDLSSSDGCRKQSS  
NKGNEAVSSSCRVRANYIGTQIPPPCH"  
4778. .4921  
/genes="T9A21.20"  
/number=1  
4922. .5022  
/gene="T9A21.20"  
/number=1  
5023. .5187  
/gene="T9A21.20"  
/number=2  
5188. .5262  
/gene="T9A21.20"  
/number=2  
5263. .5552  
/gene="T9A21.20"  
/number=3  
5553. .5637  
/gene="T9A21.20"  
/number=3  
5638. .5866  
/gene="T9A21.20"  
/number=4  
5867. .5950  
/gene="T9A21.20"  
/number=4  
5951. .6211  
/gene="T9A21.20"  
/number=5  
6969. .8087  
/gene="T9A21.30"  
join(6969. .7148,7191. .8087)  
/gene="T9A21.30"  
/note="similarity to W30DMY32F, W25DMY32F"  
/codon\_start=1  
/product="putative protein"  
/protein\_id="CAA16790.1"  
/db\_xref="GI:2832692"  
/db\_xref="TrEMBL:O49722"  
/translation="WVEPEGKFTSEERHSHYWRVLSLYVTLLAGETIATILGLRY  
YEKGKSTWLETLPSPKTKTITKKTSSFLTSLVYIGLGLVAGHCILYSFGL  
LYLPVSTSLISQLANAVSFYNSQKTPPFLNSLLVLTISTSLIVTQHEPSP  
SSTSKASKVIGTICAGSSAGTSLVSLTDYAFELKKYTKFRAILDMATPSP  
VATCVVGLFSGGWMKLSSTEMEFLQKSSYLINISTISWQACHIGSLGIEV  
SSLFNSVISTLCPLVPVPLAVVFFRDEMISGIKLVAMFLAIFGWFSVGYQHYVYNDRKE  
EDELQPSKEEBEQVDTHVQA"  
6969. .7148  
/gene="T9A21.30"  
/number=1  
7149. .7190  
intron

exon /gene="T9A21.30"  
/number=1  
7191. .8087  
/gene="T9A21.30"  
/number=2  
8612. .13959  
/gene="T9A21.50"  
join(8612. .8838,8884. .9860,10782. .11155,11210. .11915,  
12857. .13959)  
/gene="T9A21.50"  
/note="similarity to W30DMY32F, W15DMY32F  
contains EST gb:T42237"  
/codon\_start=1  
/product="putative protein"  
/protein\_id="CAA16792.1"  
/db\_xref="GI:2832694"  
/db\_xref="TrEMBL:O49724"  
/translation="MEITQVIYVNGQDASRSDVYLILPANLLFLMFSDHNNIEANLT  
QOENMTTMEISSVPSQSNKYKWLRSIYVYVYENGKSTMTGLVOLIGFPVLFL  
FRFFSQTKNPKTEADRFKFSFTILGSYIVTGLLVGANSYMSVGLLYLPVSTFSL  
ILASOLATAFYFLNSQKTPFTVNSLFLTITSSALLVNTDSVGLYVSRVKVI  
GICTIGASAGIGLLLSVLILRKVLAKQTFSTVDLVAQSLVASCVVLIGLFASG  
EWTLTSMENYKLGKVPYVMTLASIALISWQYITIGVGLIFESSVSFNSITAVGLP  
IVPVAVLVFHDKNASKIFSIILAIWGFISFYQHYLDEKCLKTSHTSPVGDPHLLP  
ABEGVNTIHSQDNLEANLIDHEVWTESSSAPQPTENYKRWLRVSIYVIFVLCQPLA  
TILGRLYYENGSTYVTVLLQLIGFPVLVLPFRFRIQPKSTDTNFSQSPSFTTLA  
SVYLCTGLLSAVYLSALATAFYFLNSQKFTPLIVSSLLLTSSALLAVNTDS  
ENSTNSRVQYVIGICTIGASAGIGLLLSIQLMFRKVTKHTSSAVTDLAIYQSLV  
ASCVVLIGLFASGWEITLPSEMRNYKLGKSYVLTLASAIISWQYITIGVGLIFESS  
SVFNSITAVGLPIPVAAVIVFHDMDASKFSIILAIQGLFSLSPYQHYLDEKCLNT  
SHTSAVGDHLPEEGHTNIQSDQNEANLIDHEETESFVPQTKRWLRVSIYAI  
FVIFCQPLATVLEGRYENGSTYVTVLLQLIGFPVLILFRFRIQPKSTDTNFS  
QSPFTTLASVLTCLGLLSAVYLSAVLGLVSTFSLILASOLATAFYFLNS  
OKFTPLIVNSLFLTSSALLAVNTDSNTNSRVQYVIGICTIGASAGIGLVL  
IQLLRKVTKHTSSAVDLANYQSLVATCVVLIGLFASGWEITLPSEMRNYKLGKVS  
YILTASAAIFWQYTVGCVGLIFESSVSFNSITAVGLPIPVVAVIVFHDMDASK  
IFSIIILAIWGLFSFYQHYLDEKCLKTCQTKPVEETQTL"  
8612. .8838  
/gene="T9A21.50"  
/number=1  
8839. .8883  
/gene="T9A21.50"  
/number=1  
8884. .9860  
/gene="T9A21.50"  
/number=2  
9861. .10781  
/gene="T9A21.50"  
/number=2  
10782. .11155  
/gene="T9A21.50"  
/number=3  
11156. .11209  
/gene="T9A21.50"  
/number=3  
11210. .11915  
/gene="T9A21.50"  
/number=4  
11916. .12856  
/gene="T9A21.50"  
/number=4  
12857. .13959  
/gene="T9A21.50"  
Query Match 100.0%; Score 1293; DB 8; Length 82697;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1293; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 GGAAGTCCTCTAGAGTGGTCATGGAATAACTCAAGTAATCTATGTCATGGAAGCAA 60  
|||||  
Db 8591 GGAAGTCCTCTAGAGTGGTCATGGAATAACTCAAGTAATCTATGTCATGGAAGCAA 8650  
Qy 61 GATGCATCTCGTAGTGTAGATTACTTGTATTCTTTTCGCTAACCTGTTGTTTGTGTTT 120

Db 8651 GATGATCTCGTAGTGTAGTACTTGTGTTCTTTTCGCTAACCTGTTGTTTGTGATGTTT 8710  
Qy 121 TCAGATCATACAACATAGAGCAAAACCTAAACAGGTCAAGAGGAAATGAATACCAACCATG 180  
Db 8711 TCAGATCATACAACATAGAGCAAAACCTAAACAGGTCAAGAGGAAATGAATACCAACCATG 8770  
Qy 181 GAATCGAATCTTCTGTCGCTACCTCAATCGAAGACTATAGAAATGCGTTCGATTTCC 240  
Db 8771 GAATCGAATCTTCTGTCGCTACCTCAATCGAAGAACTATAAGAAATGCGTTCGATTTCC 8830  
Qy 241 ATTTACGTGTTCTTGTCTGCTTGCCTGCAAGCACTTTCTACAATTTTGGGAGAGTTTAC 300  
Db 8831 ATTTACGTGTTCTTGTCTGCTTGCCTGCAAGCACTTTCTACAATTTTGGGAGAGTTTAC 8890  
Qy 301 TATGAAATGTTGGGAAGAGTACATGATGGAACACTTGTCCAACTAATCGGCTTCCT 360  
Db 8891 TATGAAATGTTGGGAAGAGTACATGATGGAACACTTGTCCAACTAATCGGCTTCCT 8950  
Qy 361 GTTCTGTTCTCTTCCGCTTCTTTTCCCAACCAAAATCCCAACCAAGAGGAGAT 420  
Db 8951 GTTCTGTTCTCTTCCGCTTCTTTTCCCAACCAAAATCCCAACCAAGAGGAGAT 9010  
Qy 421 TTCAGAAAGTCTCTTCTCTTCCATTCTTGATCAGTTTACATCGTTTACTGGACTATTA 480  
Db 9011 TTCAGAAAGTCTCTTCTCTTCCATTCTTGATCAGTTTACATCGTTTACTGGACTATTA 9070  
Qy 481 GTGCTGCTAATCTTATATATGCTCTGTTGGTTTACTATCTACCAAGTTTCTACTTTTC 540  
Db 9071 GTGCTGCTAATCTTATATGCTCTGTTGGTTTACTATCTACCAAGTTTCTACTTTTC 9130  
Qy 541 TCCCTCATCTTGGCTTCACATTTGGCTTCTACTGCTTCTTCTATATTTTCTAACTCG 600  
Db 9131 TCCCTCATCTTGGCTTCACATTTGGCTTCTACTGCTTCTTCTATATTTTCTAACTCG 9190  
Qy 601 CAGAAGTTTCAACCTTTCATTTGTGAATCTCTGTTTCTCTTACTATTTCTCTGCCCCTC 660  
Db 9191 CAGAAGTTTCAACCTTTCATTTGTGAATCTCTGTTTCTCTTACTATTTCTCTGCCCCTC 9250  
Qy 661 CTGCTGGTCAACACTGATTCGGAAGCAACAGCAAAAGTGTCTAGAGTAAATATGTGATA 720  
Db 9251 CTGCTGGTCAACACTGATTCGGAAGCAACAGCAAAAGTGTCTAGAGTAAATATGTGATA 9310  
Qy 721 GGGATAATATGATACCATTTGGTCTTCTGCTGGGATTTGGATTTGCTGCTATCCCTGGTACAA 780  
Db 9311 GGGATAATATGATACCATTTGGTCTTCTGCTGGGATTTGGATTTGCTGCTATCCCTGGTACAA 9370  
Qy 781 CTGATCCTCAGGAAGTTTAAAGAGCAAAACATTTCTCAACGGTCACTGACTTGGTCTGCT 840  
Db 9371 CTGATCCTCAGGAAGTTTAAAGAGCAAAACATTTCTCAACGGTCACTGACTTGGTCTGCT 9430  
Qy 841 TACCAATCTCTAGTTGCAAGCTGTGTGTTCTCATAGGACTTTTTCGCAAGCGGGAGTGG 900  
Db 9431 TACCAATCTCTAGTTGCAAGCTGTGTGTTCTCATAGGACTTTTTCGCAAGCGGGAGTGG 9490  
Qy 901 AAACCTTTAAACAGTGAGATGGAAACATCAAACTGGGGAAGTGCCTATGATGACT 960  
Db 9491 AAACCTTTAAACAGTGAGATGGAAACATCAAACTGGGGAAGTGCCTATGATGACT 9550  
Qy 961 TTGGCTCGATAGCTATTTCTCTGGCAAGTCTACACCATTTGGCTGCTGGGACTGATCTTT 1020  
Db 9551 TTGGCTCGATAGCTATTTCTCTGGCAAGTCTACACCATTTGGCTGCTGGGACTGATCTTT 9610  
Qy 1021 GAGTCATCTTCTGTGTTCTCCAAATCCATAAATCTGCTGGGATTCGCTATAGTTCCAGTT 1080  
Db 9611 GAGTCATCTTCTGTGTTCTCCAAATCCATAAATCTGCTGGGATTCGCTATAGTTCCAGTT 9670  
Qy 1081 CTAGCAGTGATTTGTTTCCATGATAAATGAACCGGTCAAGAGTCTTCTCCATCATTTTA 1140  
Db 9671 CTAGCAGTGATTTGTTTCCATGATAAATGAACCGGTCAAGAGTCTTCTCCATCATTTTA 9730  
Qy 1141 GCTATCTCGGGATTCATTTTCTATCTATCAGCACTACCTCGACGAAAGAGTTTGAAG 1200

Db 9731 GCTATCTGGGATTCATTTTCAATTTGTCTATCAGCACTACCTCGACGAAAGAGTTGAAG 9790  
Qy 1201 ACTAGCCACACAAGTCTCTGTAGGAGATCCTCATCTACTACTCTGTAGGAGGTTCAACA 1260  
Db 9791 ACTAGCCACACAAGTCTCTGTAGGAGATCCTCATCTACTACTCTGTAGGAGGTTCAACA 9850  
Qy 1261 AACATACATAGTGTATGATCAAAACATATTTCC 1293  
Db 9851 AACATACATAGTGTATGATCAAAACATATTTCC 9883  
RESULT 4  
ATCHRIV48  
LOCUS ATCHRIV48 194143 bp DNA linear PLN 16-MAR-2000  
DEFINITION Arabidopsis thaliana DNA chromosome 4, contig fragment No. 48.  
ACCESSION AL161548  
VERSION AL161548.2 GI:7268604  
KEYWORDS  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.  
REFERENCE 1 (bases 1 to 111084)  
MURPHY, G., RIDLEY, P., HUDSON, S., MEWES, H.W., LEMCKE, K. and MAYER, K.F.X.  
Unpublished  
REFERENCE 2 (bases 107966 to 194143)  
HILBERT, H., BRAUN, M., HOLZER, E., BRANDT, A., DUESTERHOEF, A., MEWES, H.W., LEMCKE, K. and MAYER, K.F.X.  
Unpublished  
REFERENCE 3 (bases 1 to 194143)  
EU Arabidopsis sequencing, project.  
DIRECT SUBMISSION  
TITLE Submitted (10-MAR-2000) MIPS, at the Max-Planck-Institut fuer Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail: lemcke@mips.biochem.mpg.de, mayer@mips.biochem.mpg.de  
COORDINATOR: Mike Bevan, Molecular Genetics Department, Cambridge Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK, E-mail: michael.bevan@bbsrc.ac.uk  
COMMENT Information on performance of analysis and a more detailed annotation of this entry and other sequences of chromosomes 3, 4 and 5 can be viewed at: http://www.mips.biochem.mpg.de/proj/thal/ this fragment has an overlap with ATCHRIV47 at the 5' end and an overlap with ATCHRIV49 at the 3' end.  
FEATURES  
source  
1. 194143  
/organism="Arabidopsis thaliana"  
/mol\_type="genomic DNA"  
/variety="Columbia"  
/db\_xref="taxon:3702"  
/chromosome="4"  
6377. .9293  
/gene="AT4g18120"  
join(6377. .6796, 7122. .7238, 7347. .7501, 7578. .8456, 8543. .8627, 8716. .8784, 8871. .9293)  
/gene="AT4g18120"  
/notes="similarity to Me12-like protein, Arabidopsis thaliana, gb:D86122 contains EST gb:W43784, AI998063.1"  
/codon\_start=1  
/product="putative protein"  
/protein\_id="CAB78814.1"  
/db\_xref="GI:7268605"  
/db\_xref="GOA:Q9SVV9"  
/db\_xref="TrEMBL:Q9SVW9"  
/translation="MASGLDFSGIGNMLDGDSDHPIGNMLPDEBELFSGMLDDL LSLPATLDLEDYDLFGSCGGELETDPYDSLKGFSGMFGADSNVDNVPQNI FQNVGVSIAGEHYGHPRTLFVRNINSVESELOALPEDNPSKVDNQGLVFNLP LSVNRDLNIFGVYGEIKETRETPNKRHKHFEFFDVRSDAALKALNETAGKRI KLHRSRPGGARNNMLQMPLEDDDSYSLNHFVEPLASSPIIDWRNSPIDPLQSP SKSPIFGNLSPTKNIRYPEFSMTKASVYNDOEGRRFSLDLHLFSSSYNNASHKATP QPQFSFGSVSSFGSLNSHPHVETLSGSEFLWGSPPSSAVPNVPFSSNRENHRFPYSA

QNGSLHQLHHIGSAPSGFPFRSPETSSMGSAVFRGASGNMNAQNRNIRETSSNFMKLS  
APRSQRTGNGSNLYMPAAATMWSIDDDPFGSNQOQDSNGNQADIKIQQLDLSKIMR  
GEDPRTLMIKNIPIKNTYRNMLLAAIDERNSTYDFLYLPIDFKKNKNVGYAFINNV  
PTTIALYFAFNGKWFNSKVASLAVRQGAALIAHFQNSLMNDRRCOPIV  
FQGESKYPWNLLQOHTDLKMIYILFCVYLMNLTHTCLFETRSGGEFAARASNSTG  
FLIFLSARTHIKADSQNVHVIASGLKRRKQFLG"  
6377..6796

/gene="AT4g18120"  
/number=1

6797..7121

/gene="AT4g18120"  
/number=1

7122..7238

/gene="AT4g18120"  
/number=2

7239..7346

/gene="AT4g18120"  
/number=2

7347..7501

/gene="AT4g18120"  
/number=3

7502..7577

/gene="AT4g18120"  
/number=3

7578..8456

/gene="AT4g18120"  
/number=4

8457..8542

/gene="AT4g18120"  
/number=4

8543..8627

/gene="AT4g18120"  
/number=5

8628..8715

/gene="AT4g18120"  
/number=5

8716..8784

/gene="AT4g18120"  
/number=6

8785..8870

/gene="AT4g18120"  
/number=6

8871..9293

/gene="AT4g18120"  
/number=7

9294..13230

/gene="AT4g18130"  
complement (join (9594..9837,9944..10237,10323..11130,  
11238..13230))

9594..13230

/gene="AT4g18130"  
complement (join (9594..9837,9944..10237,10323..11130,  
11238..13230))

9594..13230

/gene="AT4g18130"  
/note="Contains ATP/GTP-binding site motif A (P-loop)  
AA634-641;Phytochrome chromophore attachment site  
AA317-326

contains EST gb:T21020, AA395462"

/codon\_start=1

/product="phytochrome E"

/protein\_id="CAB78815.1"

/db\_xref="GI:7268606"

/db\_xref="GOA:P42498"

/db\_xref="Swiss-Prot:P42498"

/translation="MGFESSSSAASNMKPOKSNATQYSDAALPADPAOSIVTGKS  
FNYSKVSIPPNHVPDEHITAYLSNIQRGLVQPGCLIAVEEPSFRILGLSDNSDF  
LGLSLSPSTHSGEFDKVGKLGIDARTLFTPSSGASLSKAAASFTETSLNPNVHRS  
TTQKPFYALHRIADIGIVMDLEPAKSGDPALTAGVQSKLAVRAISRLOSLPGDI  
GALCDTVDRVQVQYDRVYQFHEDDHGEVSEIRSDLEPYLGLHYPATDIPQA  
ARFLPKONVRMI CDCAITPVKVQSEELKRLPLCLVNSTLRAPHGCHQTYMANMGSA  
SLALAIVVKGDSSKLGNLTVGHGCSPRVVPFLYACEFLMQAFGLQLOLQASQ  
LAEKAMRTQTLCDMLLRDTSVAIVTQSPGIMDLVKCDGAALYKGCWLVGVTPNE  
SQVKDLVNWLVENHGDSDGLTFTDSLVDAGYPGAI SLGDVAGVAAAFSSKDYLLWF  
RSNTASAIKMGGAHKHPKDKODAGRMPHRSFPTAFLEVAKRSRLPWEISEIDAHSRL

exon

intron

exon

intron

exon

intron

exon

intron

exon

intron

exon

intron

exon

gene

gene

CDS

LIMRSFTSSRPVLSGNGVARDANBELTSFVCEMVRVITETATAPIFGVDSSGCGINGMK  
KTAENTGLLASEAMGKSLADEIVQESRAALESLLCKALQGEESKSVMLKRFQNN  
HPDVSVCVLVNSCTSDYTENIIGVCFVGODITSEKAITDRFIRLQCDYKTIYQSL  
NPLIPIIPASDENACCSWNAAMEKLTCMSKHEVIGKMLPGEVFGVCKVKQDSLTAK  
FLISLYQIAGDNVPESSELIVFEFNKEGKYIEASLTANKSTNIEGKIVRCFFFLQILINK  
ESGLSCPBLKSAQSLNELTYVRQIKNPLRNGIRFAHKLLESSEISASORQFLETSDA  
CEKQITTIESTDLSKBEGKLEQTEFRIENILDTTISQVMIILRENSQRVEVA  
BEIKTPLNGDRVLQQLIADLRNIVNHAPPNSWVGISISPGQELSRDNGRYIHLQ  
FRMIHPGKGLPEMLSDMFETRDGWTVPDGLGLKLSRKLLQMGNGRVSFVREDERCFF  
QVDLQVKMLGVESRGTRGSSSIK"

exon

/gene="AT4g18130"

/number=1

complement (9838..9943)

/gene="AT4g18130"

/number=1

complement (9944..10237)

/gene="AT4g18130"

/number=2

complement (10238..10322)

/gene="AT4g18130"

/number=2

complement (10323..11130)

/gene="AT4g18130"

/number=3

complement (11131..11237)

/gene="AT4g18130"

/number=3

complement (11238..13230)

/gene="AT4g18130"

/number=4

14593..16312

/gene="AT4g18140"

complement (join (14593..14751,14921..15070,15338..15639,  
15915..16055,16141..16312))

/gene="AT4g18140"

complement (join (14593..14751,14921..15070,15338..15639,  
15915..16055,16141..16312))

/gene="AT4g18140"

/note="similarity to HYA22 protein, Homo sapiens,  
PIR2:JCS707

contains EST gb:R65235"

/codon\_start=1

/product="putative protein"

/protein\_id="CAB78816.1"

/db\_xref="GI:7268607"

/db\_xref="TrEMBL:Q9SVV8"

/translation="MKSISKISLREKRVLGVCKKPKHDKVNSCFDPKISERMELLVLT  
ATQCNNEQNI GLDSTRDVVTNEPQDLVEDSSKNASNNMTIFSPVLDDLYSET  
DRPESVLIDAAEQYMLLPYLEDRSASSDDVKSDNDRINRSRKNLEAAENHTEAQT  
EDFDPQIFLRNQPELADVFNYPVDMQPPRDSPKRKAVTFVLINMLLDILOPDGKFVS  
ORFYRDSCLSDGIYTKDLTVLGLDLAKVALVDNCPQYVRLQINNGIFIKSWYDDPTD  
DGLTLPLFLETLDANDVRPIAKRFNGKE"

complement (14593..14751)

/gene="AT4g18140"

/number=1

complement (14752..14920)

/gene="AT4g18140"

/number=1

complement (14921..15070)

/gene="AT4g18140"

/number=2

complement (15071..15337)

/gene="AT4g18140"

/number=2

complement (15338..15639)

/gene="AT4g18140"

Query Match 100.0%; Score 1293; DB 8; Length 194143;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1293; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAAGTCCTCTTAGAGTGGTCATGGAATAACTCAAGTAATCTATGTCAATGTAAGCAA 60

Db 36978 GGAAGTCTCTTAGAGTGGTCATGGAAATAACTCAAGTAATCTATGTCAATGGTGAAGCA 37037  
Qy 61 GATGATCTCGTAGTGTAGATTACTTGTGATCTTTTCGCTAACTGTTGTTTTCGATGTTT 120  
Db 37038 GATGATCTCGTAGTGTAGATTACTTGTGATCTTTTCGCTAACTGTTGTTTTCGATGTTT 37097  
Qy 121 TCAGATCATACAACATAGAACCAACCTTAACAGGTACAGGAGGAAAATGAATACCACCATG 180  
Db 37098 TCAGATCATACAACATAGAACCAACCTTAACAGGTACAGGAGGAAAATGAATACCACCATG 37157  
Qy 181 GAAATCGAATCTTCGTCGGTCACTCAATCGAAGAACTATAAGAAATGSCCTTCGTAATTC 240  
Db 37158 GAAATCGAATCTTCGTCGGTCACTCAATCGAAGAACTATAAGAAATGSCCTTCGTAATTC 37217  
Qy 241 ATTTACGTTCTTTTGTCTCTTGTGCAAGCACCTTTCTACAATTTTGGGCAGAGTTTAC 300  
Db 37218 ATTTACGTTCTTTTGTCTCTTGTGCAAGCACCTTTCTACAATTTTGGGCAGAGTTTAC 37277  
Qy 301 TATGAAATGTTGGGAAGATACATGATGGGAACACTTGTCCAACTAATCGGCTTCCT 360  
Db 37278 TATGAAATGTTGGGAAGATACATGATGGGAACACTTGTCCAACTAATCGGCTTCCT 37337  
Qy 361 GTTCGTTTCTCTTCGCTCTCTTTCCCAACCAAAATCCAAACCAAGAGAGAGAT 420  
Db 37338 GTTCGTTTCTCTTCGCTCTCTTTCCCAACCAAAATCCAAACCAAGAGAGAGAT 37397  
Qy 421 TTCGAAAGTTCTCTTCCTCCACCATCTTCGATCAGTTTACATCGTTACTGGACTATTA 480  
Db 37398 TTCGAAAGTTCTCTTCCTCCACCATCTTCGATCAGTTTACATCGTTACTGGACTATTA 37457  
Qy 481 GTGTCGTCACTCTTATATGTCCTCTGTTGGTTTACTATTAATACAGTTTCTACTTTC 540  
Db 37458 GTGTCGTCACTCTTATATGTCCTCTGTTGGTTTACTATTAATACAGTTTCTACTTTC 37517  
Qy 541 TCCCTCATCTTGGCTCACAATTTGGCTTCTACTGCTTCTCTCATATTTTCTAAACTCG 600  
Db 37518 TCCCTCATCTTGGCTCACAATTTGGCTTCTACTGCTTCTCTCATATTTTCTAAACTCG 37577  
Qy 601 CAGAAGTTTACACCTTTTCATTTGTAATTTCTGTTTTCCTTACTATTTTCTCGCCCTC 660  
Db 37578 CAGAAGTTTACACCTTTTCATTTGTAATTTCTGTTTTCCTTACTATTTTCTCGCCCTC 37637  
Qy 661 CTCGTGTCACACTGATTCGGAAACACACGAAAGTGTCTAGAGTAATATGTGATA 720  
Db 37638 CTCGTGTCACACTGATTCGGAAACACACGAAAGTGTCTAGAGTAATATGTGATA 37697  
Qy 721 GGGATAATATGTACCATTTGGTCTTCTGCTGGGATTTGGATTTGCTATCCCTCGTACAA 780  
Db 37698 GGGATAATATGTACCATTTGGTCTTCTGCTGGGATTTGGATTTGCTATCCCTGGTACAA 37757  
Qy 781 CTGATCTCAGGAAGTTTAAAGAACCAACATTTCTCAACGGTCACCTGACTTGGTCGCT 840  
Db 37758 CTGATCTCAGGAAGTTTAAAGAACCAACATTTCTCAACGGTCACCTGACTTGGTCGCT 37817  
Qy 841 TACCAATCTCTAGTTGCAAGCTGTGTGTTCTCATAGACATTTTCGCAAGGGGGAGTGG 900  
Db 37818 TACCAATCTCTAGTTGCAAGCTGTGTGTTCTCATAGACATTTTCGCAAGGGGGAGTGG 37877  
Qy 901 AAACTTTAAACAAGTGAGATGGAAACCTACAAACTGGGGAAGTGCCATACGTTATGACT 960  
Db 37878 AAACTTTAAACAAGTGAGATGGAAACCTACAAACTGGGGAAGTGCCATACGTTATGACT 37937  
Qy 961 TTGGCCTCGATAGCTATTTCTTCGCAAGTCTACACCAATTTGGCGTCGTGGGACTGATCTTT 1020  
Db 37938 TTGGCCTCGATAGCTATTTCTTCGCAAGTCTACACCAATTTGGCGTCGTGGGACTGATCTTT 37997  
Qy 1021 GAGTCATCTTCTGTGTTCTTCCAAATCCCAATCTGCTGTGGGATTTGCCTATAGTTCCAGTT 1080  
Db 37998 GAGTCATCTTCTGTGTTCTTCCAAATCCCAATCTGCTGTGGGATTTGCCTATAGTTCCAGTT 38057  
Qy 1081 GTAGCAGTGATTTTTCATCATGATAAATGAACCGTCAAGAGATCTTCTCCATCATTTTAA 1140

Db 38058 GTAGCAGTGATTTGTTTCCATGATAAAATGAACGGTCAAGAGATCTTCTCCATCATTTTA 38117  
Qy 1141 GCTATCTGGGGATTCATTTTCATTTGTTCTATCAGCACTACCTCGACGAAAAGATTGAAG 1200  
Db 38118 GCTATCTGGGGATTCATTTTCATTTGTTCTATCAGCACTACCTCGACGAAAAGATTGAAG 38177  
Qy 1201 ACTAGCCACACAAGTCTCTGTAGGAGATCCTCATCTACTACTCTGTAGGAGGTCACACA 1260  
Db 38178 ACTAGCCACACAAGTCTCTGTAGGAGATCCTCATCTACTACTCTGTAGGAGGTCACACA 38237  
Qy 1261 AACATACATAGTGTATGATCAAAACATATTTCC 1293  
Db 38238 AACATACATAGTGTATGATCAAAACATATTTCC 38270  
  
RESULT 5  
AX507517 3387 bp DNA linear PAT 27-SEP-2002  
LOCUS  
DEFINITION Sequence 2212 from Patent WO0216655.  
ACCESSION AX507517  
VERSION AX507517.1 GI:23388754  
KEYWORDS  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.  
1  
REFERENCE  
AUTHORS Harper, J.F., Kreps, J., Wang, X. and Zhu, T.  
TITLE Stresses-regulated genes of plants, and transgenic plants containing same, and methods of use  
JOURNAL Patent: WO 0216655-A 2212 28-FEB-2002;  
The Scripps Research Institute (US) ; Syngenta Participations AG (CH)  
FEATURES  
source 1. .3387  
/organism="Arabidopsis thaliana"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:3702"  
  
ORIGIN  
Query Match 89.1%; Score 1152; DB 6; Length 3387;  
Best Local Similarity 96.4%; Pred. No. 7.6e-286;  
Matches 1207; Conservative 0; Mismatches 0; Indels 45; Gaps 1;  
Qy 22 ATGGAATAAATCACTCAAGTAATCTATGTCAATGGTAAGCAAGATGCATCTCGTAGTGTAGAT 81  
Db 1 ATGGAATAAATCACTCAAGTAATCTATGTCAATGGTAAGCAAGATGCATCTCGTAGTGTAGAT 60  
Qy 82 TACTTGATTTCTTTTCGCTAACCTGTTGTTTTCGATGTTTTCAGATCATACACATAGAA 141  
Db 61 TACTTGATTTCTTTTCGCTAACCTGTTGTTTTCGATGTTTTCAGATCATACACATAGAA 120  
Qy 142 GCAAACTCAACAGCTCAGGAGGAAATGAATACCAACCATGGAATTCGAATCTTCGTCGCTA 201  
Db 121 GCAAACTCAACAGCTCAGGAGGAAATGAATACCAACCATGGAATTCGAATCTTCGTCGCTA 180  
Qy 202 CCTCAATCGAAGAACTATAAGAAATGCGCTTCGTAATTTCCATTTACGTTGTTCTTTGTCCTT 261  
Db 181 CCTCAATCGAAGAACTATAAGAAATGCGCTTCGTAATTTCCATTTTACGT----- 227  
Qy 262 GCTTGCCAAGCAGCTTTCTACAAATTTGGGCAGAGTTTACTATGAATAATGGTGGGAAGAGT 321  
Db 228 -----AGTTTACTATGAATAATGGTGGGAAGAGT 255  
Qy 322 ACATGGATGGGAACACCTTGTCCAACTAATCGGCTTCCTCTGTTCTCTTCCTCCGCTTC 381  
Db 256 ACATGGATGGGAACACCTTGTCCAACTAATCGGCTTCCTCTGTTCTCTTCCTCCGCTTC 315  
Qy 382 TTTTCCCAACCAAAATCCAAACCAAGAGAGAGATTTTCAGAAAGTTCTCTTCCTTC 441  
Db 316 TTTTCCCAACCAAAATCCCAACCAAGAGAGATTTTCAGAAAGTTCTCTTCCTTC 375



```
QY 922 GAAAACTACAACTGGGGAAGTGCATACGTTATGACTTTGGCTCGATAGCTATTTC 981
Db 856 GAAAACTACAACTGGGGAAGTGCATACGTTATGACTTTGGCTCGATAGCTATTTC 915
QY 982 TGGCAAGTCTACACCATTTGGCGTGGTGGGAGTGTATGAGTCACTTTCTGTTCTCC 1041
Db 916 TGGCAAGTCTACACCATTTGGCGTGGTGGGAGTGTATGAGTCACTTTCTGTTCTCC 975
QY 1042 AATTCCATAAGTGTGGGATTCCTATAGTTCAGTGTGAGTGTGAGTGTGTTTCCAT 1101
Db 976 AATTCCATAAGTGTGGGATTCCTATAGTTCAGTGTGAGTGTGAGTGTGTTTCCAT 1035
QY 1102 GATAAAATGAACGGCTCAAGATCTTCTCCATCAATTTAGTATCTGCGGATTCATTCA 1161
Db 1036 GATAAAATGAACGGCTCAAGATCTTCTCCATCAATTTAGTATCTGCGGATTCATTCA 1095
QY 1162 TTTGTCTATCAGCTACTCTCGACGAAAGTGTGAAGTGTAGCTAGCCACACAAGTCTGTA 1221
Db 1096 TTTGTCTATCAGCTACTCTCGACGAAAGTGTGAAGTGTAGCTAGCCACACAAGTCTGTA 1155
QY 1222 GGAGATCCTCATCTACTACTCTGCTGAGGAAGTGCACACACATACATAGTG 1273
Db 1156 GGAGATCCTCATCTACTACTCTGCTGAGGAAGTGCACACACATACATAGTG 1207

RESULT 7
AX033548 LOCUS AX033548 1194 bp DNA linear PAT 21-SEP-2000
DEFINITION Sequence 5 from Patent WO0049152.
ACCESSION AX033548
VERSION AX033548.1 GI:10280292
KEYWORDS
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1
REFERENCE
AUTHORS Andre,B., Buurkle,L., Frommer,W.B. and Gillissen,B.
TITLE Nucleic acids that code for a nucleobase transporter
JOURNAL Patent: WO 0049152-A 5 24-AUG-2000;
ANDRE BRUNO (BE) ; BUERKLE LUKAS (DE) ; FROMMER WOLF B (DE) ;
GILLISSEN BERND (DE)
FEATURES
source location/Qualifiers
1..1194
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"

ORIGIN
Query Match 63.2%; Score 816.6; DB 6; Length 1194;
Best Local Similarity 81.8%; Pred. No. 2.2e-199;
Matches 971; Conservative 0; Mismatches 204; Indels 12; Gaps 2;
QY 107 TGTGTTTTCATGTTTTCAGATCATACACATAGAGCAAACTAACAGTTCAGGAGGAAA 166
Db 20 TGAGTGTAAATTTTTCAGTGACCAAGAACTTAGAAGCAAACTTATAGATCATGAGGTGG 79
QY 167 TGAATACCAACCATGGAATCGAATCTTCGTCGCTACCTCAATCGAAGAACTATAGAAAT 226
Db 80 TAACT-----GAATCATCATCATCAGCTGTGCTTCAAAACCGAGAACTATAAAGGT 130
QY 227 GGCTTCGTATTTCCATTTACGTGTTCTTTGTCCTTGTGTCGCAAGCACTTTCTCAATTT 286
Db 131 GGCTTCGTGTCCTCATATACGTAATCTTTGTCTCTTTTGGCCAGCCACTAGTACAAATTC 190
QY 287 TGGGCGAGAGTTTACTATCAAAATGTTGGGAGAGTACATGATGGAGCACTTGTCCAAAC 346
Db 191 TGGGTAGATTGTACTATCAAAATGAGGAATAGCACAATATGTGGTACACTTCTTCAC 250
QY 347 TAATCGGCTTCCCTGTTCTGTTCTTCTCCGCTTCTTTTCCCAACCAAAATCCCAAC 406
Db 251 TCATTGGCTTCCCTGTTCTGTTCTTCTTCTCCGCTTCTTTCTCGAATCAGGCAACCCAAT 310
```

```
QY 407 CAACAGAGCAGATTTTCAGAAAGTCTCTCTCCCTTACCATTCTTGGATCAGTTTACATCG 466
Db 311 CAACAGATACAAATTTTCAGTCAGTCCCTTCCCTTACCACCCCTTGCATCGGTTTACTTGT 370
QY 467 TTACTGGACTATTAGTGTCTGTCTAACTCTTATATGTCCTCTGTGTTGTTTACTACTATTAC 526
Db 371 GCAGTGGACTGTAGTGTCTGCGCTTATGCTTATTTGTCGAGTAGGGTTCCTCTACTTAC 430
QY 527 CAGTTTCTACTTTCTCCCTCATCTTGGCCCTCACAAATTTGGCCCTTCACTGGCTTCTTCTCAT 586
Db 431 CAGTCTCTACTTTCTCCCTCATCTTGGCCCTCACAGTTTGGCCCTTCACTGGCTTCTTCTCAT 490
QY 587 ATTTTCTAAACTCGCAGAAGTTCACACTTTTCAATTTGGAATTTCTCTGTTTCTCCTTACTA 646
Db 491 ATTTTCTTAACTCGCAAAAGTTTCACTCTTTGATGATGTCAGTTCTTTGCTTCTCCTCACTG 550
QY 647 TTTTCTCTGCGCCCTCTCGTGGTCAACACATGATTCGGAACACACAGCAAAAAGTGTCTAGAG 706
Db 551 TATCCTCTGCTCTCTTGTGTCACACATGATTCAGAAAACTCAACTAATGATATCTAGAG 610
QY 707 TAAAAATATGATAGGGAATAATATGACCAATTTGGTGTCTTCTGCTGGGATTTGGATTTGCTGC 766
Db 611 TACAGTATGTGATCGGGTTCATATGTACCAATCGGTGCTTCCGCTGGGATTTGGACTGTATAC 670
QY 767 TATCCCTGTGTACAACTGATCCTCAGGAAGCTTTTAAAGAAAGCAAAACATTCTCAACGCTCA 826
Db 671 TATCTGTGATACAAATGCTCTTTCAGGAAGTTTTCAGAAAGCATACATCTCTCAGCAGTCA 730
QY 827 CTGACTTTGGTCTGCTTTACCAATCTCTAGTTTCAAGCTGTGTGGTTTCTCATAGGACTTTTTCG 886
Db 731 CGGACTTTGGCCATTTACCACTCTAGTTTGGAGTTGTGTAGTTCTCATAGGACTTTTTCG 790
QY 887 CAAGCGGGAGTGGAAACCTTTTAAAGTGAATGGAATACTACAACTCGGGGAAAGTGC 946
Db 791 CAAGTGGAGAGTGGGAACTTTTGCCTAGTGAATGGAATACTACAACTCGGGGAAAGTGT 850
QY 947 CATACGTTATGACTTTTGGCCCTCGATAGTATTTTCTCGCAAGTCTACACCATTTGGCCTCG 1006
Db 851 CATATGTTTGTAGCTTTTAGCTTCGCGAGCTATTTTCTGGCAAGTCTACACTCTTGGTCTTG 910
QY 1007 TGGGACTGTACTTTTGTAGTCTATCTTCTGTGTTCTTCCAAATTCATAACTGTGTGGGATTCG 1066
Db 911 TGGGATTTGATCTTCGAGTCACTCTCTGTGTTCTTCCAAATTCATAACAGCTGTGGGATTCG 970
QY 1067 CTATAGTTCCAGTTGTAGCAGTGAATTTTTCATGATAAAATGAAGCGCTCAAGATCT 1126
Db 971 CTATAGTTCCAGTTGTGGCGAGTGTAGTTTTCATGATAGATGGAGCGCATCAAAAATCT 1030
QY 1127 TCTCCATCATTTTAGCTATCTGGGATTCATTTTCATTTGTCTATCAGCACTTACCTCGACG 1186
Db 1031 TCTCCATATTTTAGCTATCTGGGCTTCTTTTCATTTCTGCTATCAGCACTTACCTCGACG 1090
QY 1187 AAAAGAAAGTTGAAGACTAGCCACACAGTCTCTGTAGGAGATCTCTATCTACTACCTGCTG 1246
Db 1091 AAAAGAAAGTTGAATACTACTAGCCACACAGTCTGTAGGAGATCTTCTAT---CTACCTGTTG 1147
QY 1247 AGGAAGTGCACAAACATACATAGTGTATGATCAAAACATATTTTTC 1293
Db 1148 AGGAAGTGCACAAACATACAAAGTGTGTATCAAAAGCATATTTTTC 1194
```

```
RESULT 8
AY096558 LOCUS AY096558 1165 bp mRNA linear PLN 18-SEP-2002
DEFINITION Arabidopsis thaliana putative protein (At4g18200) mRNA, complete cds.
ACCESSION AY096558
VERSION AY096558.1 GI:20465496
KEYWORDS FLI CDNA..
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
```



Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

#### REFERENCE AUTHORS

1 (bases 1 to 1165)  
Yamada, K., Banh, J., Chan, M.M., Chang, C.H., Chang, E., Dale, J.M.,  
Deng, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L.,  
Tang, C.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Bowser, L.,  
Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J.,  
Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B.,  
Lin, J., Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J.,  
Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A.,  
Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.  
Arabidopsis Open Reading Frame (ORF) Clones  
Unpublished

#### TITLE JOURNAL REFERENCE AUTHORS

2 (bases 1 to 1165)  
Yamada, K., Banh, J., Chan, M.M., Chang, C.H., Chang, E., Dale, J.M.,  
Deng, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L.,  
Tang, C.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Bowser, L.,  
Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J.,  
Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B.,  
Lin, J., Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J.,  
Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A.,  
Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.  
Direct Submission

#### TITLE JOURNAL COMMENT

Submitted (16-APR-2002) Plant Gene Expression Center, 800 Buchanan  
Street, Albany, CA 94710, USA  
The RIKEN Genomic Sciences Center (GSC) members carried out the  
collection and clustering of RAPL cDNAs (RAPL cDNA: 'RIKEN  
Arabidopsis Full-Length cDNA'): Seki, M., Narusaka, M., Ishida, J.,  
Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,  
Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PGECC (SSP) Consortium members constructed and  
sequenced the pUNI (ORF) clones using the RAPL cDNAs: Yamada, K.,  
Banh, J., Chan, M.M., Chang, C.H., Chang, E., Dale, J.M., Deng, J.M.,  
Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C.,  
Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Bowser, L., Chen, H.,  
Cheuk, R., Jones, T., Karlin-Neumann, G., Kim, C., Lam, B., Lin, J.,  
Meyers, M.C., Miranda, M., Nguyen, M., Palm, C.J., Shinn, P.,  
Southwick, A., Davis, R.W., Ecker, J.R. and Theologis, A.

Yamada, K. (SSP/PGECC) and Seki, M. (RIKEN GSC) contributed equally  
to this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP  
/PGECC) contributed equally to this work as PIs.

Annotation is based on the January 2002 version of the Arabidopsis  
genome submitted to GenBank.

#### FEATURES source

Location/Qualifiers  
1..1165  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
/chromosome="4"  
/clone="U18528"  
/ecotype="Columbia"  
/note="This clone is in pUNI 51."

#### Gene CDS

1..1165  
/gene="At4g18200"  
1..1134  
/gene="At4g18200"  
/codon\_start=1  
/evidence=experimental  
/product="putative protein"  
/protein\_id="AA020208.1"  
/db\_xref="GI:20465497"  
/translation="MGISQVHYCNGDQNLNLDHEETSESVPTQKNCWLRVSI  
YALFVFCPLATVLRGLYENGKSTVYVTLQLIGFVPLILFRFRIRQPKSTDI  
NFSQSPFTLASVLCGLLSAVAYLSAGLVLPVSTFSLIASQLAFAPFSYF  
LNSQKFTPLIVNSLFLTVSSALLVNTSDSENTVRSYVYVIGFICIGASGIGLV  
LSLIQLFRKVFRTKTSASVLDLVAVTCVVLIGLFAGSEWRVLPSEMRNYKLG  
KVSYLTLTASAAIFQWYVGVCGVGLIFSSSVFSNSITAVGLPIVPPVAVIVFDKMD  
ASKIFSIILAIWGLFSLFVQYHLDKELKTCQTKPVESETQL"  
1135..1165  
/gene="At4g18200"

#### ORIGIN

Query Match 59.9%; Score 775; DB 8; Length 1165;  
Best Local Similarity 83.8%; Pred No. 1.2e-188;  
Matches 877; Conservative 0; Mismatches 170; Indels 0; Gaps 0;  
QY 180 GGAAATCGAATCTTCGTCCGTACTCAATCGAAGAACTATAAGAAATGCGTTCTGTAATTC 239  
Db 69 GGAAATCGAATCATTTTCAGTACCTCAACAGGAAGAACTGAAGAGGTGGCTCGGTGCTC 128  
QY 240 CATTTACGTGTTCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGT 299  
Db 129 CATATACGCAATCTTTGTCTATCTTCTGCAACCACTTGTCTGAGTCTGCGTAGACTGTA 188  
QY 300 CTATGAAATGCTGGCAAGAGTACATGATGGGAACACTTGTCCAACTAATCGGCTTCCC 359  
Db 189 CTATGAAATGAGGAGGAAAGACATATGTGTATACATCTTCTCAACTCATTTGGCTTCCC 248  
QY 360 TGTTCGTGTTCTTTCGCGTCTTTCCTTCCAAACCAAAATCCAAACCAAGAGCAGA 419  
Db 249 TGTACTGATCTCTTTCGCGTCTTTCCTTTCGAAATCAGGCAACCAAAATCAACAGATACAA 308  
QY 420 TTTGAGAAAGTCTCTCTTCCCTTCCATCTTGTGGATCAGTTTACATCGTTTACTGGACTATT 479  
Db 309 TTTGAGTCACTCCCTTCTTCCACCACTTGTGATCGGTTTACTTGTGCACTGGACTGCT 368  
QY 480 AGTGTCTGTAACCTCTATATGTCCTCTGTTGGTTTACTATCTTACAGATTTCTACTTT 539  
Db 369 AGTGTCTGTTATGCTTATTTGTGTCAGTTGGTTGCTTTTACTTACCAAGTCTCTACTTT 428  
QY 540 TTCCCTCATCTTGGCCTTCAAAATGGCCTTCACTGCTCTTCTTCTCATATTTTCTAAACTC 599  
Db 429 CTCCCTCATCTTGGCCTTCAAGTTGGCTTCACTGCTCTTCTCATATTTCTTAACTC 488  
QY 600 GCAGAGTTTCAACCTTTTCATTTGGAATCTCTGTTTCTTCTTCTTCTTCTTCTTCTGCTC 659  
Db 489 GCAGAGTTTCACTCTCTTGTGATGTCATTTCTTGTGTTCTTCTTCTTCTTCTTCTTCTGCTC 548  
QY 660 CTCTGTTGTCACACTGATTCGGAACACACAGCAAAAGTGTCTAGAGTAAATATGTGAT 719  
Db 549 CTCTGTTGTCACACTGATTCGGAACACACAAATGTATCTAGAGTAAATATGTGAT 608  
QY 720 AGGGAATAATGTATACCTTGTGCTTCTGCTGGGATTTGGATTTGCTGCTATCCCTGGTACA 779  
Db 609 TGGTGTCTATCTGTACCTTGTGCTTCTGCTGGGATTTGGATTTGCTGCTATCTCTGATACA 668  
QY 780 ACTGATCTCAGGAAGTTTAAAGAACAAACATCTCAACGGTCACTGACTTGGTGGC 839  
Db 669 ACTGCTCTCAGGAAGTTTAAAGAACAAACATCTCAACGGTCACTGACTTGGTGGC 728  
QY 840 TTACCAATCTCTAGTTGCAAGTGTGCTTCTCATAGGACTTTTTCGCAAGGGGAGTG 899  
Db 729 TTACCAAGTCTCTAGTTGCAAGTGTGCTTCTCATAGGACTTTTTCGCAAGGGGAGTG 788  
QY 900 GAAATCTTAACTAGATGGAATCTCAAACTGGGGAAGTSCCATACGTTATGAC 959  
Db 789 GAGAACTCTGCAAGTGAATGAGAACTCAAACTGGGGAAGTGTATATATCTTAC 848  
QY 960 TTTGGCTCGATAGCTATTTCTGCGCAAGTCTACACCATTTGGCGTGTGGAGTCACTCTT 1019  
Db 849 TTTGGCTCGATAGCTATTTCTGCGCAAGTCTACACCATTTGGCGTGTGGAGTCACTCTT 908  
QY 1020 TGAGTCATCTTCTGTTTCTCAATTCCTGCTGGGATTTGGCTATAGTTTCCAGT 1079  
Db 909 CGAGTCTTCTGTTTCTCAATTCCTGCTGGGATTTGGCTATAGTTTCCAGT 968  
QY 1080 TGTAGCAGTATGTTTTCATGATTAATGAACGCGTCAAGAGATCTTCTCCATCATTTT 1139  
Db 969 TGTAGCAGTATGTTTTCATGATTAATGAACGCGTCAAGAGATCTTCTCCATCATTTT 1028  
QY 1140 AGCTATCTGGGATTCATTTCTATTTGCTATCAGCACTACCTCGACGAAAGAGTGA 1199  
Db 1029 AGCTATCTGGGCTTCTTCTTCTATTTGCTATCAGCACTACCTCGACGAAAGAGTGA 1088

#### 3' UTR



```
Db      800  TTACCAAGTCTCTAGTGTGCAACTTGTGTGGTACTCATAGGACTGTGTTTGCAGGTGAGAGTG 859
Qy      900  GAAAACTTTTAAACAAGTGAGATGAAAACTACAAACTGGGAAAGTGGCCATACGTTATGAC 959
Db      860  GAGAACTCTGCCAAGTGAGATGAGAACTACAAACTGGGAAAGTGTCTATATCTTGAC 919
Qy      960  TTTGGGCTCGATAGCTATTTCCTGGCAAGTCTACACCATTTGGCGTCTGGGACTGATCTT 1019
Db      920  TTTGGGCTCAGCAGCTATTTCCTGGCAAGTATACACTGTGGTGTGGGATTAATCTT 979
Qy      1020  TGAGTCATCTTCTGCTTCTCCAAATTCCTCAATTCCTCAATTCCTCAATTCCTCAATTCCT 1079
Db      980  CGAGTCTTCTTCTGCTTCTCCAAATTCCTCAATTCCTCAATTCCTCAATTCCTCAATTCCT 1039
Qy      1080  TGTAGCAGTGATTGTTTTCATGATAAAATGAACGGTCAACAGATCTTCTCCATCATTTT 1139
Db      1040  TGTAGCAGTGATGATTTCATGATGAAGATGATGATGATGATGATGATGATGATGATGATGAT 1099
Qy      1140  AGCTATCTGGGATTCATTTCTATCTATCAGCACTACCTCGACGAAAGATGAA 1199
Db      1100  AGCTATCTGGGCTTCTTCTTCTATCTATCAGCACTACCTCGACGAAAGATGAA 1159
Qy      1200  GACTAGCCACAAAGTCTGTAGGAGA 1226
Db      1160  GACTTCCGACAAACCTGTGTAGGA 1186
```

```
RESULT 10
AF370622
LOCUS   AF370622.1 GI:13877726
DEFINITION Arabidopsis thaliana putative protein (T9A21.60) mRNA, complete
cds.
ACCESSION AF370622
VERSION   AF370622.1
KEYWORDS  FLI CDNA.
SOURCE    Arabidopsis thaliana (thale cress)
ORGANISM  Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 1106)
Lam,B., Southwick,A., Karlin-Neumann,G., Nguyen,M., Miranda,M.,
Palm,C.J., Bowser,J., Jones,T., Banh,J., Chen,H., Cheuk,R.,
Chung,M.K., Kim,C., Lin,J., Liu,S.X., Pham,P.K., Sakano,H.,
Shinn,P., Yamada,K., Ecker,J., Theologis,A. and Davis,R.W.
Direct Submission
Submitted (17-APR-2001) DNA Sequencing and Technology Center,
Stanford University, 855 California Avenue, Palo Alto, CA 94304,
USA
e-mail for correspondence: arab@sequence.stanford.edu
This clone was isolated by RT-PCR.
```

```
FEATURES
source
    1..1106
        /organism="Arabidopsis thaliana"
        /mol_type="mRNA"
        /db_xref="taxon:3702"
        /chromosome="IV"
        /clone="C531"
        /ecotype="Columbia"
    1..1106
        /gene="T9A21.60"
        657..1106
        /gene="T9A21.60"
        /codon_start=1
        /product="putative protein"
        /protein_id="AAK43941.1"
        /db_xref="GI:13877727"
        /translation="MDMIYVSLVASCVSWGLFASSEWKTLSSEMDNYKHGKVSIM
NLVATVWQVFSIGTGLIFELSLFSNAISVLGLPVVPIILAVLIFDKNLKLAVIS
MILAIWGFYSYIQYQLDDNLUKNHEIITTESPDPEABESTWQSK"
ORIGIN
```

```
Query Match      35.1%; Score 454.4; DB 8; Length 1106;
Best Local Similarity 65.2%; Pred. No. 4.9e-106;
Matches 687; Conservative 0; Mismatches 361; Indels 6; Gaps 1;

Qy      199  GTACCTCAATCGAAGAACTATAAGAAATGGCTTGTATTTCCATTTACGTGTTCTTTGTC 258
Db      36  GTATCTCACTCTAAACACATACAAACGGTGGCTCAGGGTACTTTATATACATTTCTTTGTC 95
Qy      259  CTTGCTTGGCAAGCACTTTCTACAATTTTGGGACAGTTTACTATGAAAATGGTGGGAAG 318
Db      96  ATTTGAGGCCAAACAGTTGCTACAAATTTTGGCAGAGTATATATGACAAACGAGGAAC 155
Qy      319  AGTACATGATGGGAACACTTGTCCAATTAATCGCTTCCCTGTTCTGTGTTCTCTTCCGC 378
Db      156  AGTAAATGGCTAGCAACGGTAGTTCAACTTGTGTGCTTCCCTGCTACTTCCATATTAT 215
Qy      379  TTCCTTTTCCCAACCAAAATCCCAACCAACAGAGAGATTTGAGAAAGTTCCTTTC 438
Db      216  ATCTTTGTCAATTTTAAACACATGCAACAACTGATAGAGATG-----GAAAAAGAACCTCA 269
Qy      439  TTCACCATTTCTTGGATCAGTTTACATCGTTACTGGACTATTAGTGTCTGCTAACTCTTAT 498
Db      270  CCTAGGAACCGTGTATGTTTACGTAGTGTGGACTTCTTGTAGGAGCAGATTGCTAT 329
Qy      499  ATGTCTCTGTGTTGTTTACTATATACCAAGTTCCTACTTTCTCCCTCATCTTGGCCTCA 558
Db      330  CTGTACTCCATTGGACTTCTTTACTTACCCTGTTTCTACCTATTCCTGATCTGTGCATCT 389
Qy      559  CAAATTGGGCTTCACTGCCTTCTTCTCATATTTTCTAAACTCGCAAGATTTCAACCTTTC 618
Db      390  CAGTTAGGCTTCAATGCTTCTTCTTATTTTCTTAACTCAAAAACTTACCCCTATC 449
Qy      619  ATTGTGAATCTCTGTTCTCTTACTATTTCTCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 678
Db      450  ATTTTAAATCTCTTTTCTCTTAACTATATCTTCCACCTACTTGCATTTCAATATAGAG 509
Qy      679  TCGGAAACACACAGCAAAAGTGTCTAGAGTAAATATGTATAGGATATATATGACCAATT 738
Db      510  GAGACAGACTCCACAAAAGTTTACAAAAGAGAGATATGTCAAAGGTTTTCATATGCACGTT 569
Qy      739  GGTGCTTCTGCTGGGATTTGGATGCTATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 798
Db      570  GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 629
Qy      799  TTAAGAGAGCAAAACATTTCTCAACGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 858
Db      630  CTAAGAGAGCAAAATTTCTCAGAGTATGATATGATATGATATGATATGATATGATATGATATGAT 689
Qy      859  AGCTGTGTGTTCTCATAGGACTTTTCGCAAGCGGGAGTGGAAACTTTTAAACAAGTGAG 918
Db      690  AGTTGTGTTAGCGTGTGGGCTTTTTCGCTAGCAGTGAAGTGGAAACTTTTGAAGCAGTGAA 749
Qy      919  ATGGAATACTACAACTGGGAAAGTGCATACGTTATGACTTTTGGCCTCGATAGCTATT 978
Db      750  ATGGATACTACAAACATGGGAAGGTATCTACATTTATGAACCTAGTGTGGACAGCTGTT 809
Qy      979  TCCTGCGAAGTCTACACCATTTGGCGTGTGGGACTGATCTTTGAGTCACTTCTCTGTGTC 1038
Db      810  ACCTGGCAGGTATTTCTCCATCGTGGCAGAGACTGATCTTTCGAGCTCTCTCTCTATTC 869
Qy      1039  TCCAAATCCCAACTGTGTGGGATTTGCTATAGTTCAGTTTGTAGCAGTGAATGTTTTC 1098
Db      870  TCAATGCAATAAGCGTTTGGGACTCCCGAGTGGTTCCTATCTTGGCTGTAATCATTTTC 929
Qy      1099  CATGATAAAATGAACGCTCAAGATCTTCTCCATCATTTTACGTATCTCTGGGATTCATT 1158
Db      930  CATGACAAAATGAATGGGTAAAGGTGATTTCTATGATCCTAGCTATTTTGGGGTTTCACT 989
Qy      1159  TCATTTGCTATCAGCACTTCTCGACGAAAGATTTGAAGACTAGCCACACAGTCTCT 1218
Db      990  TCCTATGCTACCAACAATATCTTGTATGACAAAACCTTGAAGAAAATCATGAAATCACA 1049
Qy      1219  GTAGGAGATCCTCATCTACTACTCTGCTGAGGAAG 1252
```



Db 345 AGCTTCACCGCTTCTCTCTTATTACTCAACTCACAACAACTTACTCTCTATCATTTT 404  
QY 624 GAATTCCTGTTTCTCTCTTACTATTTCTCTGCTCCCTCGTGGTCAACACTGATTCGGA 683  
Db 405 GAATTCCTTCTCTCTCTCTATATCTTCCACCTCTCTTGCATTTAATACGAGGAATC 464  
QY 684 AAACACAGCAAAAGTGTCTAGAGTAAATATGTGATGAGGATATATGTACCAATTGGTGC 743  
Db 465 AGATTCACAAAAGTTTACAAAAGGAGATGTGTCAAAAGTTTCTGATGCAACCGTTGGTGC 524  
QY 744 TTCTGCTGGATTTGGATTTGCTGCTATCCCTGCTGATCAACTGATCTCAGAAAGTTTAAA 803  
Db 525 ATCTGCTGGTTGGTCTACTCTTATCCCTTACACAGCTAGCTTTTCGTAAGATTTTAAA 584  
QY 804 GAAGCAAACTTCTCAACGGTCACTGACTTGGTCTGTATCAACTCTCTAGTTGCAAGCTG 863  
Db 585 GAAGCAAACTTCTCAGAAAGTTATAAATATGATATCACTACATGAGTCTAGTGCCAGTTG 644  
QY 864 TGTGGTTCTCATAGGACTTTTTCGAAGCGGGAGTGGAACACTTTAAACAGTGAGATGA 923  
Db 645 TGTAGCTGCTGGTGGGCTTTTGTCTAGTAGCGAGTGGAAACCTTTTGACGAGTGAATGA 704  
QY 924 AACTACAACTGGGAGAGTGCATACGTTATGACTTTGGCTCGATAGCTATTTCTCTG 983  
Db 705 AACTACAACTGGGAGAGTATCCTATGTATGAACCTAGTGTGACAGCTGTATACCTG 764  
QY 984 GCAAGTCTACACATTTGGGCTCGTGGGACTGTATTTGAGTCACTCTCTGTCTCTCAA 1043  
Db 765 GCAGGTATTTCTCATCGTTGCACAGACTGATCTTCGAGCTTCTCTCTCTCTCTCAA 824  
QY 1044 TTCCATACTGCTGCTGGGATTCCTATAGTTTCAGTTGTAGCAGTGAATTTTCCATGA 1103  
Db 825 TGCATAAAGCGCTTTGGGACTCCCGCTGCTTCTATCTCTGCTGCTCATCATTTTCCATGA 884  
QY 1104 TAAATGAAGCGGTCAAGATCTTCTCATCATTTTAGCTATCTGGGATTTCAATTCATT 1163  
Db 885 CAANAATGAAGCGCTTAAAGGTTGATTTCTATGATTTCTAGCTATTTGGGTTTCTGATCTTA 944  
QY 1164 TGTCTATGAGCTACTCTCGACGAAAGAGTTTGAAGACTAGCCACACAAGTCTGTAGG 1223  
Db 945 TGTCTACCAACATATCTTTGATGAACAACTTTGAAGAAAGTATGAATACCAACAC 1004  
QY 1224 AGATCTCTATCTACTACTGCTGTGAGGAAGG 1253  
Db 1005 AGAATCCCTGACCGACCAAGACGAGG 1034

RESULT 13  
BX664674/c

LOCUS Arabidopsis thaliana transposon insertion STS AT\_5.3478, sequence tagged site.  
DEFINITION BX664674

ACCESSION BX664674.1 GI:37665166  
VERSION STS; STS, sequence tagged site.  
KEYWORDS Arabidopsis thaliana (thale cress)  
SOURCE Arabidopsis thaliana

ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

## REFERENCE

1 Clarke,J.H., Bowles,B., Carter,J., Hart,D., McCullagh,B., Walsh,S., Langham,S., LeGrys,C., Jones,J.D.G. and Bevan,M.

Unpublished

2 (bases 1 to 721)

Clarke,J.H.

Direct Submission

Submitted (14-OCT-2003) Clarke J.H., John Innes Centre, Colney

Lane, Norwich, NR4 7UJ, UK

JOURNAL AT denotes an activation tag dissociation transposon within a single line, Et an enhancer trap dissociation transposon, Gr a gene trap dissociation transposon, MT a mis-expression enhancer trap dissociation transposon, SM a defective suppressor mutator

transposon. \_3 denotes a sequence derived from the 3' end of the transposon, \_5 denotes a sequence derived from the 5' end of the transposon BBRC GARNET, ATIS project  
On-line seed stock requests: [http://nasc.nott.ac.uk/NASC\\_stock\\_code: N10051](http://nasc.nott.ac.uk/NASC_stock_code: N10051).

## FEATURES

source Location/Qualifiers  
1..721  
/organism="Arabidopsis thaliana"  
/mol\_type="genomic DNA"  
/variety="Landsberg erecta NASC stock code NW20"  
/cultivar="Th113xTh25"  
/db\_xref="taxon:3702"  
/clone="AL021713"  
1..721  
/standard\_name="AT\_5.3478"

## STS

## ORIGIN

Query Match 33.2%; Score 429.6; DB 11; Length 721;  
Best Local Similarity 87.1%; Pred. No. 1.3e-99;  
Matches 484; Conservative 0; Mismatches 99; Indels 3; Gaps 1;

QY 738 TGGTCTTCTCTGCTGGGATTGGATTCGCTGCTATCCCTGGTACAACTGATCCTCAGGAAGCT 797  
Db 709 TCGGGCTTCCGCTGGGATTGGACTGTTACTATCTCTGATACAAATGCTCTTCAGGAAGT 650  
QY 798 TTAAAGAAAGCAAAATCTCAACGGTCACTGACTTGGTGGCTTACCATCTCTAGTTGC 857  
Db 649 TTTCAGGAAGCATACATCTCAGCAGTCAAGGACTTGGCCATTTACCAGTCTCTAGTTGC 590  
QY 858 AAGCTGTGTGTTCTCATAGGACTTTTCGAAGCGGGAGTGGAAAACTTTAAACAGTGA 917  
Db 589 GAGTTGTGTAGTTCTCATAGGACTTTTTCGAAGTGGAGAGTGGAAAACTTTGCCAAGTGA 530  
QY 918 GATGAAACACTACAACTGGGGAAGTGGCATACTGATTTAGCTTTGGCTCGATAGCTAT 977  
Db 529 GATGAAACACTACAACTGGGGAAGTGTCTATGTTTGTAGCTTTAGCTTCGGCAGCTAT 470  
QY 978 TTCTTGGCAAGTCTACACATTTGGGCTCGTGGGACTGATCTTTGAGTCACTCTCTGTGT 1037  
Db 469 TTCTTGGCAAGTCTACACTCTTGGTCTTGGGATTTGATCTTCGAGTCACTCTCTGTGT 410  
QY 1038 CTCCAAATCCATAACTGCTGTGGGATTCGCTATAGTTCCAGTTGTAGCAGTGAATGTTT 1097  
Db 409 CTCCAAATCCATAACAGCTGTGGGATTCGCTATAGTTCCAGTTGGGCAAGTATAGTTT 350  
QY 1098 CCATGATAAATGAACGCTCAAGATCTTCTCCATCATTTTAGCTATCTGGGATTCAT 1157  
Db 349 CCATGATAAATGAACGCTCAAAATCTTCTCCATTTTGTAGCTATCTGGGCTTCTCT 290  
QY 1158 TTCAATTTGCTATCAGCACTACCTCGACGAAAGAGTTGAAGACTAGCCACACAAGTCC 1217  
Db 289 TTCAATTTGCTATCAGCACTACCTCGACGAAAGAGTTGAATATCTAGCCACACAAGTCC 230  
QY 1218 TGTAGGAGATCCTCATCTACTACTGCTGTGAGGAAGGTCAACAAACATACATAGTGTATG 1277  
Db 229 TGTAGGAGATCTTCAT---CTACCTGTTGAGGAAGGTCAACAAACATACAAAGTGTGTG 173  
QY 1278 ATCAAAACATATTTCC 1293  
Db 172 ATCAAAAGCATATTTCC 157

## RESULT 14

AX412868

LOCUS

DEFINITION

Accession

Version

Keywords

Source

Organism

AX412868 Sequence 632 from Patent WO022675.  
AX412868 AX412868.1 GI:21445326  
Arabidopsis thaliana (thale cress)  
Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

AX412868 Sequence 632 from Patent WO022675.  
AX412868 AX412868.1 GI:21445326

Arabidopsis thaliana (thale cress)  
Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Arabidopsis thaliana (thale cress)  
Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Arabidopsis thaliana (thale cress)  
Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Arabidopsis thaliana (thale cress)  
Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Arabidopsis thaliana (thale cress)  
Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Arabidopsis thaliana (thale cress)  
Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Arabidopsis thaliana (thale cress)  
Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Arabidopsis thaliana (thale cress)  
Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Arabidopsis thaliana (thale cress)  
Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Arabidopsis thaliana (thale cress)  
Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

1  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
Patent: WO 022675-A 632 21-MAR-2002;  
Syngenta Participations AG (CH) ; UNIVERSITY OF NORTH CAROLINA AT  
CHAPEL HILL (US) ; Glazebrook, Jan (US) ; Wang, Xun (US) ; Dangl,  
Jeffrey L. (US) ; Eulgem, Thomas (US)  
FEATURES  
source  
1..1047  
/organism="Arabidopsis thaliana"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:3702"

ORIGIN  
Query Match 31.7%; Score 410; DB 6; Length 1047;  
Best Local Similarity 65.0%; Pred. No. 1.4e-94;  
Matches 624; Conservative 0; Mismatches 330; Indels 6; Gaps 1;

QY 293 GAGTTACTATGAAATGGTGGAGAGTACATGATGGGAACACTTGTCCAATATCG 352  
DB 71 GAGTATACTATGACACCGGAGAAACAGATAATGGCTAGCAACGGTAGTTCAACTTGTG 130  
QY 353 GCTTCCCTGTTCTGTTCTCTTCCGCTTCTTCCAAACCAAAATCCCAACCAACAG 412  
DB 131 GCTTCCCTGTTCTGTTCTCTTCCGCTTCTTCCAAACCAAAATCCCAACCAACAG 412  
QY 413 AAGCAGATTTTCAGAAAGTTCTTCTTCCCTTCCAAACCAAAATCCCAACCAACAG 472  
DB 191 GAGATGGAAAGAACTTCACTTATATATATATATATATATATATATATATATATAT 190  
QY 473 GACTATATAGTCTGCTAATCTTATATATATATATATATATATATATATATATATAT 532  
DB 245 GACTTCTTGTAGGAGCAGATGCTATCTGATCTGATCTGATCTGATCTGATCTGAT 304  
QY 533 CTACTTTCTCCCTCATCTTGGCTTCCAAATGGCTTCACTGCTTCTTCTCATATTTTC 592  
DB 305 CTACTTATCTCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGAT 364  
QY 593 TAACTCGCAGAGTTTCACCTTTTCAATTTGGAATTTCTGTTTCTCTTACTATTTCT 652  
DB 365 TTAATCTCAAAAATTTACCCCTATCTTAAATTTCTTCTTCTTCTTCTTCTTCTTCT 424  
QY 653 CTGCTCTCTGTTGTCACACTGATTCGGAACACAGCAAAAGTCTAGAGTAAT 712  
DB 425 CCACCTTCTGCTTCAATATGAGGAGCAGACTCCAAAGAGTTTACAAAGGAGAGT 484  
QY 713 ATGTGATAGGATATATGATCTTGGCTTCTGCTGGATTTGGATTTGCTATCCC 772  
DB 485 ATGTCAAAGTTTCATATGACCGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 544  
QY 773 TGGTACAACTGATCCTCAGGAAGTTTAAAGAGCAAACTTCTCAACGCTCACTGACT 832  
DB 545 TACAACAGCTAGCTTTCTAAAGCTCTTAAAGAGCAAACTTCTCAGAGTTATGATA 604  
QY 833 TGGTGGCTTACCAATCTAGTTGCAAGCTGTGTTGTTCTCATAGGACTTTTCCAAAG 892  
DB 605 TGATAATCTAGCTGAGTCTAGTGCCAGTTGTGTAGCGTGGTGGGCTTTTCTAGCA 664  
QY 893 GGGAGTGAAGAACTTTAAACAGTGGATGGAAACTACAACTGGGAAAGTGCCATACG 952  
DB 724 GTGAGTGAAGAACTTTGAGGAGTGAATGGATTAATTAACATGGAAGATCTCTTACA 712  
QY 953 TTATGACTTTGGCTCGATAGTATTTCTGGCAAGTCTACCAATTTGGGCTGGGAC 1012  
DB 784 TTATGAACCTAGTGTGGACAGCTGTACCTGGCAGGATTTCTCCATCGGTGGCACAGG 784  
QY 1013 TGATCTTTGAGTCACTCTCTGTTTCTCCAAATTCATTAATCTGCTGGGATTCCTATAG 1072  
DB 785 TGATCTTTGAGTCT 844  
QY 1073 TTCCAGTTGTAGCAGTGATTTGTTTCCATGATAAAATGAACCGCTCAAGATCTTCTCCA 1132

DB 845 TTCTATCTTGGCTGTAATCATTTTCCATGCAAAATGAATGGTTAAAGGTGATTCTA 904  
QY 1133 TCATTTTGTAGTCTATCTGGGATTCATTTTCTATCAGCACTACTCTGCAAGAAAGA 1192  
DB 905 TGATCTCTAGCTATTTGGGGTTTCACTTCTCTATGCTTACCAACAATATCTTGTACAAA 964  
QY 1193 AGTTGAAGACTAGCCACACAGTCTCTAGAGATCTCTATCTACTACTGCTGAGGAG 1252  
DB 965 ACTTGAAGAAATATCATGAATCACAACAGAAATCCCTGAGCCCAAGACGAGAG 1024

RESULT 15  
AX033549  
LOCUS  
DEFINITION  
Sequence 6 from Patent WO0049152.  
ACCESSION  
AX033549  
VERSION  
AX033549.1 GI:10280293  
KEYWORDS  
Arabidopsis thaliana (thale cress)  
ORGANISM  
Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

1  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
Andre, B., Buerkle, L., Frommer, W.B. and Gillissen, B.  
Nucleic acids that code for a nucleobase transporter  
Patent: WO 0049152-A 6 24-AUG-2000;  
ANDRE BRUNO (BE) ; BUERKLE LUKAS (DE) ; FROMMER WOLF B (DE) ;  
GILLISSEN BERND (DE)  
FEATURES  
Location/Qualifiers  
source  
1..1081  
/organism="Arabidopsis thaliana"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:3702"

ORIGIN  
Query Match 31.7%; Score 410; DB 6; Length 1081;  
Best Local Similarity 65.0%; Pred. No. 1.4e-94;  
Matches 624; Conservative 0; Mismatches 330; Indels 6; Gaps 1;

QY 293 GAGTTACTATGAAATGGTGGAGAGTACATGATGGGAACACTTGTCCAATATCG 352  
DB 89 GAGTATACTATGACACCGGAGAAACAGATAATGGCTAGCAACGGTAGTTCAACTTGTG 148  
QY 353 GCTTCCCTGTTTCTTCTTCCGCTTCTTCCAAACCAAAATCCCAACCAACAG 412  
DB 149 GCTTCTCTGCTACTTCCATATATATATATATATATATATATATATATATATATAT 208  
QY 413 AAGCAGATTTTCAGAAAGTTCTTCTTCCCTTCCAAACCAAAATCCCAACCAACAG 472  
DB 209 GAGATGGAAAGAACTTCACTTACCTTCCCTTCCAAACCAAAATCCCAACCAACAG 262  
QY 473 GACTATTAGTGTCTGCTAATCTTATATGTCCTCTGTTGGTTTACTATATCTTACCAGTTT 532  
DB 263 GACTTCTTGTAGGAGCAGATTGCTATCTGCTTCCATTTGACTTCTTACTTACCGGTTT 322  
QY 533 CTACTTTTCTCCCTCACTTCTGGCTCAAAATTTGGGCTTCTGCTTCTTCTCATATTTTC 592  
DB 323 CTACTTATTCCTGATCTGTGCATCTCAGTTAGCTTCAATGCTTCTTCTTCTTATTTTC 382  
QY 593 TAACTCGGAGAGTTTCACACCTTTCATTTGTAATCTCTGTTTCTTCTTCTTACTTTCT 652  
DB 383 TTAACACAAAAAATTTACCCCTTATCATTTTAAATTTCTTCTTCTTCTTAACTATATCT 442  
QY 653 CTGCTCTCTCTGCTGGTCAACACTGATTCGGAACCAACAGCAAAAGTCTTAGAGTAAAT 712  
DB 443 CCACCTTCTTGTCAATTAATGAGGAGACAGCTCCCAAAAGTTTACAAAGGAGAGT 502  
QY 713 ATGTGATAGGATTAATGATGACATTTGGTGTCTTCTGCTGGGATTTGGATTCGCTATCC 772  
DB 503 ATGTCAAAGGTTTCATATGACCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 562

Qy	773	TGGTACAACTGATCCTCAGGAGGTTTAAAGAGCAAAATCTCTCAACGGTCACGTGACT	832
Db	563	TACAACAGCTAGCCCTTTCTAAAGAGTCTTAAGAGCAAAATTTCTCAGAAGTTATGGATA	622
Qy	833	TGGTCGCTTACCAATCTCTAGTTGCAAGCTGTGTCTCTCATAGGACTTTTCGCAAGCG	892
Db	623	TGATAATCTACGTGATCTAGTGGCCAGTTGTGTAGCGTGGTGGGGCTTTTGTCTAGCA	682
Qy	893	GGGAGTGGAAAACTTTTAAACAAGTGAGATGGAAAACTTACAAACTTGGGGAAAGTGCCATAG	952
Db	683	GTGAGTGGAAAACTTTGAGCAGTGAATGGATAACTTACAAACATGGGAAGGTATCCTACA	742
Qy	953	TTATGACTTTGGCTCGATAGCTATTTCTTGGCAAGTCTACACCATTTGGCGTGGTGGAC	1012
Db	743	TTATGAACCTAGTGTGGACAGCTGTACCTGGCAGGTATTTCTCCATCGGTGGCACAGGAC	802
Qy	1013	TGATCTTTGAGTCATCTTCTGTCTCTCCAAATTTCCATAACTGTGTGGGATTGCCCTATAG	1072
Db	803	TGATCTTCGAGCTCTCTCTATTTCTCAATGCAATAAGCGTTTGGGACTCCCAAGTGG	862
Qy	1073	TTCCAGTTGTAGCAGTGATTTTTCATGATAAAAATGAACGCGTCAAAAGATCTTCTCCA	1132
Db	863	TTCCATATCTTGGCTGTAAATCAITTTCCATGACAAAATGAATGGGTTAAAGGTGATTTCTA	922
Qy	1133	TCATTTTAGCTATCTGGGATTCATTTCTATTTGTCTATCAGCACTACCTCGACGAAAAA	1192
Db	923	TGATCCTAGCTATTTGGGGTTTTCATCTTCTATGTCTACCAACAATATCTTGATGACAAA	982
Qy	1193	AGTTGAAGACTAGCCACACAAGTCCGTAGGAGATCCTCATCTACTACCTGCTGAGGAG	1252
Db	983	ACTTGAAGAAAAATCATGAATCACAAACAGATCCCTGACCCACCAGAAAGCAGAAG	1042

Search completed: November 1, 2004, 15:44:38  
Job time : 5735.91 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 1, 2004, 06:20:04 ; Search time 593.039 seconds  
(without alignments)  
9900.026 Million cell updates/sec

Title: US-09-913-767-3

Perfect score: 1145

Sequence: 1 ctgtttcattgtgtgtgtg.....caagtttaggaagatcca 1145

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3413475 seqs, 2563800928 residues

Total number of hits satisfying chosen parameters: 6826950

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications\_NA.\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq.\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq.\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*
- 13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq.\*
- 16: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq.\*
- 17: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq.\*
- 18: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*
- 19: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*
- 20: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*
- 21: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	354	30.9	1294	16	US-10-424-599-2524
2	318.4	27.8	3387	9	US-09-938-842A-2212
3	318.4	27.8	3387	11	US-09-938-842A-2212
4	281.8	24.6	1513	17	US-10-437-963-82061
5	280.4	24.5	2214	16	US-10-424-599-102519
6	243.2	21.2	83698	17	US-10-416-898-9
7	217	19.0	997	16	US-10-424-599-75161
8	197.2	17.2	2381	17	US-10-437-963-43860
9	168.6	14.7	728	17	US-10-767-701-9000
10	163	14.2	747	17	US-10-767-701-6806
11	154.6	13.5	592	9	US-09-770-152-44
12	130.6	11.4	2605	17	US-10-437-963-83900

13	128.4	11.2	792	16	US-10-424-599-9213	Sequence 9213, Ap
14	127	11.1	1779	16	US-10-425-114-28473	Sequence 28473, A
15	125.4	11.0	1386	16	US-10-425-114-12485	Sequence 12485, A
16	125	10.9	1432	17	US-10-437-963-1897	Sequence 1897, Ap
17	111.8	9.8	522	16	US-10-424-599-114931	Sequence 114931, A
18	105.2	9.2	327	16	US-10-424-599-121152	Sequence 121152, A
19	104.4	9.1	332	11	US-09-732-627A-1196	Sequence 1196, Ap
20	97	8.5	1146	16	US-10-425-114-10516	Sequence 10516, A
21	97	8.5	1233	16	US-10-424-599-30274	Sequence 30274, A
22	87.4	7.6	1373	16	US-10-424-599-31364	Sequence 31364, A
23	86.6	7.6	650	17	US-10-767-701-4045	Sequence 4045, Ap
24	84	7.3	1333	16	US-10-425-114-10574	Sequence 10574, A
25	83.4	7.3	671	17	US-10-767-701-5986	Sequence 5986, Ap
26	78.4	6.8	1355	16	US-10-424-599-31363	Sequence 31363, A
27	76.2	6.7	774	9	US-09-770-445-895	Sequence 895, App
28	75	6.6	308	16	US-10-424-599-121805	Sequence 121805, A
29	53.2	4.6	1417	16	US-10-424-599-97043	Sequence 97043, A
30	52.8	4.6	386	9	US-09-770-791-85	Sequence 85, Appl
31	52.8	4.6	716	16	US-10-424-599-61310	Sequence 61310, A
32	49	4.3	13573	15	US-10-311-455-1842	Sequence 1842, Ap
33	47.4	4.1	613	17	US-10-767-701-4026	Sequence 4026, Ap
34	45	3.9	3090	16	US-10-654-416-5	Sequence 5, Appli
35	44.6	3.9	17213	15	US-10-311-455-1456	Sequence 1456, Ap
36	44	3.8	15373	15	US-10-311-455-440	Sequence 440, App
37	43	3.8	6145	16	US-10-221-714A-269	Sequence 269, App
38	42.6	3.7	3081	16	US-10-654-416-9	Sequence 9, Appli
39	42.6	3.7	46732	17	US-10-741-601-5752	Sequence 5752, Ap
40	42.4	3.7	492	16	US-10-424-599-92714	Sequence 92714, A
41	42.2	3.7	631	17	US-10-767-701-5979	Sequence 5979, Ap
42	42	3.7	402	16	US-10-131-837-8426	Sequence 8426, Ap
43	42	3.7	3673778	15	US-10-312-841-1	Sequence 1, Appli
44	41.6	3.6	476	10	US-09-814-353-4700	Sequence 4700, Ap
45	41.6	3.6	476	10	US-09-814-353-10999	Sequence 10999, A

ALIGNMENTS

RESULT 1  
US-10-424-599-2524  
; Sequence 2524, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 2524  
; LENGTH: 1294  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_102285C.1  
US-10-424-599-2524

Query Match	30.9%	Score 354;	DB 16;	Length 1294;
Best Local Similarity	59.6%	Pred. No. 2.9e-89;		
Matches 619;	Conservative 0;	Mismatches 410;	Indels 9;	Gaps 1;
Qy	33	GAATTTTCAACAGAGAAGTCAACAAGTACTCTTGGAGGTTAAGAGTGTCTCTCTA	92	
Db	94	GAATGAATCAACATGAGTAAAGAAAAGATATTACCGTTGGCTCAGATATCCATCCA	153	
Qy	93	TGTCACTCTCTCTTAGCTGGAGAGACAATAGGCACTCTCTTAGTAGACTTACTAGA	152	
Db	154	TTCTTCACTTGTGTGGTGTGCGGATACGAGCAATCTCTCTTGAAGATTGTACTATGA	213	



Db 3031 AGTGGAGAGTGAGAACTCTGCCAAGTAGAGAGAGAACTACAAACTGGGAAAGTGTC 3090  
Qy 808 TACATTTTGAATAAACAATCGGTTCAACGATATCATGGCAAGCTTGTGATTGGAAGTGTT 867  
Db 3091 TATATCTTGACCTTGGCCCTCAGCAGCTATTTCTGCAAGATATACACTGTTGGTGTG 3150  
Qy 868 GGTGATGATATGAAGTTTCAATGCTTTTTCGAATGTCATGAAGCACTTTGTTTACCA 927  
Db 3151 GGAATTAATCTCGAGTCTTCTTCTGTTCTCCTCAATTCATTAATCTGCTGGGACTGCT 3210  
Qy 928 GTTGGCCGTTCTTCTGCTGCTCTTCTTCTCGTGTGATGATGAGTGAATCAAGTTGTT 987  
Db 3211 ATAGTTCAGTTGTAGCAGTAGTATTTTCCATGATGAATGATGATCAAAAGATTTTC 3270  
Qy 988 GCAATGTTTGGCCATCTGGGATTTGTTCTTATGTTTATCAGCATTTATGTCATGAT 1047  
Db 3271 TCCATCATTTTAGTATCTGGGGCTTCCCTTTCATTTGCTATCAGCACTACCTCGAGAA 3330  
Qy 1048 AGAAG 1053  
Db 3331 AAGAAG 3336

## RESULT 3

US-09-938-842A-2212  
; Sequence 2212, Application US/09938842A  
; Publication No. US20040009476A9  
; GENERAL INFORMATION:  
; APPLICANT: Harper, Jeff  
; APPLICANT: Kreps, Joel  
; APPLICANT: Wang, Xun  
; APPLICANT: Zhu, Tong  
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
; FILE REFERENCE: SCRIPT300-3  
; CURRENT APPLICATION NUMBER: US/09/938,842A  
; CURRENT FILING DATE: 2001-08-24  
; PRIOR APPLICATION NUMBER: US 60/227,866  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: US 60/264,647  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/300,111  
; PRIOR FILING DATE: 2001-06-22  
; NUMBER OF SEQ ID NOS: 5379  
; SEQ ID NO 2212  
; LENGTH: 3387  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-09-938-842A-2212

Query Match 27.8%; Score 318.4; DB 11; Length 3387;  
Best Local Similarity 57.8%; Pred. No. 6.4e-79;  
Matches 593; Conservative 0; Mismatches 421; Indels 12; Gaps 1;

Qy 28 GAAGGAAATTTTCAACGAGAGAGAGTCAAGTACTTTGAGAGTTAAGAGTGCT 87  
Db 2323 GAAACTGAATCAATTTTCAGTACTCTCAACGAAAGACTGTAAGAGTGGTCTCGTCTCC 2382  
Qy 88 CTCTATGTCACCTCTCTTACTCGAGAGACAATAGCCACTCTCTTAGGTAGACTTAC 147  
Db 2393 ATATACGAATCTTTGTCTATCTTTCGCCAACCACTTGTCTACAGTTCTCGGTAGACTGTAC 2442  
Qy 148 TACGAAAAGGCGGTAAAGACATCGGCTCGAAACCTTGGTTCAGCTTGTAGGGTTTCT 207  
Db 2443 TATGAAATGAGGAAAGACATATGTGTAACACTTCTTCAACTCATTTGGCTTCCCT 2502  
Qy 208 TTAACCCCTTCTTGTCTATTTACTTAAAGCCTGAGCCGTCAAGACTTAAACCACTTACC 267  
Db 2503 GTACTGATCTCTCCGCTCTTCTTCTCGAATCAGGCAACCAATCAACAGATACAAAT 2562  
Qy 258 AAAAAAACTACTTCTCTTCTTGTGACATATCTTTAGTGTATATTTGGACTTGGCTTGT 327  
Db 2563 TTCAGTCAGTCCCTTCTCTTACCACCCTTGCATCGGTTTACTTGTGCACTGGACTGCTA 2622

Qy 328 GTTGCTGACATTTGATTTTGTACTCATTTGGGCTACTTTACCTTCTCTCTCAACTTTC 387  
Db 2623 GTGCTGCTATGCTTATTTGTCTGAGTGGTGGTGTCTTACTTACCAGTCTCTACTTTC 2682  
Qy 388 TCTTTGATCTCTGCTGCAATTTGGCTTTTAAAGCCGCTTCTCTTACTTCTTAAACTCA 447  
Db 2683 TCCCTCATCTTGGCCCTCAGTGGCTTCACTGCTTTTCTCATATTTTCCCTTAACTCG 2742  
Qy 448 CAAAAAATCACACATTTATATCAATTCACATTTGTTCTCTTAAACATATCTTCTACATTT 507  
Db 2743 CAAAAGTTCACTCTTGTATAGTCAATTTCTTGTCTCTTACGGTTTCTCTGCCCC 2802  
Qy 508 CTGTTATCCAAACATGAACCAATCTCCCTCTTCTACTTCAAAAGTCCGCGACCAAGTCC 567  
Db 2803 CTCGGTCAACACTGATTCAGAA-----AACACAACAATGATCTAGAGTA 2850  
Qy 568 AAGTATGATTTGATATACATCTGCGCGGTGCTAGCTCAGCTGGTTATTTCTCTGGTGT 627  
Db 2851 CAATATGATTTGGGTTTCACTGTACCATTTGGTGTCTCCGCTGGGATTTGGATTTGTA 2910  
Qy 628 TCTTTAACAGATTACGCGTTGAAAGATTTCTAAGAAATACACATTTCAAGGCTATTTTA 687  
Db 2911 TCTCTGATACACTGCTCTTCAGGAAAGTTTTCAGAAAGCATACATCTCAGCAGTCTTG 2970  
Qy 688 GACATGGCCACATATCCCTCTATGCTAGTACTTGTGTAGTTTGTGTAGGACTTTTGTGGA 747  
Db 2971 GACTTGGCCAAATACCAGTCTCTAGTTGCAACTTGTGTGTACTCATAGGACTGTTTGA 3030  
Qy 748 AGTGTGGTGGAAAAAGCTGAGTACAGAAATGGAAGAGTTTCAACTAGGGAAGAGTCA 807  
Db 3031 AGTGAGAGTGGAGAACTCTGCAAGTGAAGATGAGAACTACAAACTGGGAAAGTGTCA 3090  
Qy 808 TACATTTTGAATAACATCGGTTCAACGATATCATGCGCAAGCTTGTGTTGGAAGTGT 867  
Db 3091 TATATCTTGACTTTTGGCCCTCAGCAGCTATTTCTTGGCAAGTATACACTGTTGTTGTG 3150  
Qy 868 GGTTTGATTTATCGAAGTTTTCATCGCTTTTTCCTTCCAAATGTCATAAGCACTCTTTGT 927  
Db 3151 GGATTAATCTTCGAGTCTTCTCTGTTCTTCCAAATTCATTAATCTGCTGGAGTCTCT 3210  
Qy 928 GTTGTGCTGTTCTTGTGTTGCTTCTTCCGTTGATGAGATGAGTGGAAATCAAGTTGTT 987  
Db 3211 ATAGTTCAGTTGTAGCAGTATAGTTTTCATGATGAAGATGGATGATCAAAAGATTTTC 3270  
Qy 988 GCAATGTTTGGCCATCTGGGATTTGTTCTTATGTTTATGCTTATCAGCATTTATCTCA 1047  
Db 3271 TCCATCATTTTAGTATCTGCGGCTTCTTTCATTTGTCTATCAGCACTACCTCGACGAA 3330  
Qy 1048 AGAAG 1053  
Db 3331 AAGAAG 3336

## RESULT 4

US-10-437-963-82061/c  
; Sequence 82061, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966

; SEQ ID NO 82061  
; LENGTH: 1513  
; TYPE: DNA  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_81525C.1  
US-10-437-963-82061

Query Match 24.6%; Score 281.8; DB 17; Length 1513;  
Best Local Similarity 57.0%; Pred. No. 9.4e-69;  
Matches 564; Conservative 0; Mismatches 407; Indels 18; Gaps 2;

```
Qy 70 TGGAGGTTAAGAGTCTCTCTATGTCACTCTCTCTAGCTGAGAGACAATAGCCACT 129
Db 1258 TGGTGGCTGATGGTGGAGTTAAACATGTTCTCTCATTTGCCGTGACGCGCATCTACA 1199

Qy 130 CTCTTAGTAGACTTTACTACGAAAAAGCGGTAAAAAGCACATGGCTCGAAACCTTGGTT 189
Db 1198 CTCTCTGGAGGTTCTACTACAAATCAAGCGGCAATAGCAAGTGATGTCCACATTCGTC 1139

Qy 190 CAGCTTGAGGGTTTCCTTTAAACCTTCCTTGTCTATTATTACTTTAAAGCCGTGACCGTCC 249
Db 1138 CAAACTGCTGGCTTCCGATTTGTTGATTCGCCCTATTTCTTTCCATTCAAAGACATCT 1079

Qy 250 AAGACTAAACCATTAACCAAAAAA-----CTACTTCTTCTCTTTGACACTATCTTTA 303
Db 1078 TCTACACAAACAGTCACTAGTAGTCTGCGCCCTACAATTTCTATCCCAAAATTACTCTG 1019

Qy 304 GTGTATATTAGACTTGGCTTCTGTTGCTGGACATTTGATTTTGTACTCATTTGGGCTA 363
Db 1018 ATATATGTTGCTTGGGCTCATCATTTGCTGCAGACGACTTGATGTATCTCTATGGCTA 959

Qy 364 CTCTACCTTCCTGTCCTCAACTTTCTCTTGTATCTCTGCTGCGCAATTTGGCTTTTAAAGCC 423
Db 958 CTATATCTTCGGGCTCAACATATTCGCTCATCTGTGTAGTCAGCTTGGCTTCAATGCT 899

Qy 424 GTCTTCTTACTTCTTAACTCAAAAAATCACACATTTATATCTCAATTCATTTGTT 483
Db 998 GTCTTCTCATATTTCTCAATGCTCAAAAAATTCACCCCTCTGATTTTCAATTCGTTAGTC 839

Qy 484 CTCTTAAACCATATCTTCTACACTTCTTGTATCAACATGAACCAAGATCTCCCTCTTCT 543
Db 838 CTCTTACGTTTCTGCTTCACTCTTGGAGTGTGATGAAGTTCTCAGGGAATCTACTAGT 779

Qy 544 ACTTCAAAGTCGCGACGCAAGTCCAAGTATGTGATGTGATACATCTGCGCGGTGCGTAGC 603
Db 778 ATATCA-----CATGGGAAGTACATTTTGGGTTTCTCTGTGACACTAGGGGCA 731

Qy 604 TCAGCTGGTTATCTCTGGTCTTCTTTTAACAGATTACGGTTTCGAAAGATTTCTAAG 663
Db 730 TCAGCTACATACTCGCTCATTTCTCCCTGATGCAAGTCAATTTGAAAGGTTATTAAAG 671

Qy 664 AAATACACATTTCAAGGCTATTTTAGACATGGCCACATATCCGTCTATGGTAGCTATTGT 723
Db 670 AGGAGAGACTTCTCAGTTGTGTGAAATGATGAGATATATACAGCTCTCGTGGCAACATG 611

Qy 724 GTAGTTGTGGTGAAGCTTTTGGAAAGTGGTGGGTTGGAAAGAGCTGAGTACAGAAATGAA 783
Db 610 GCTTCTCTTGTGGTTATTTGCAAGTGGTGAATGGATGACTTTTACAAGGAGAGATGCAT 551

Qy 784 GAGTTTCAACTAGGGAAGCTCATACATTTTGTGATAACATCGGTTCAAAGATATCATGG 843
Db 550 GCAATTCAGTCTGGGAAGCTGTCATATGTAATGACACTGCTGTGGAGCGGCTATATCTTGG 491

Qy 844 CAAAGCTGTTTGTGTAAGTGTGGTTGATTTATCGAAAGTTTCATCGCTTTTTTCCAAAT 903
Db 490 CAGGTAGCATCAGTTGGAGTGTGGATTTGATCTTTGTGGTGTCTATCGCTGTTTCAAT 431

Qy 904 GTCATAGACATCTTTGTTTACAGTTGTGCGCTGTTCTTGTGTGTCTTCTCCGCTGAT 963
Db 430 GTGATAAGCACCCCTAGCTCTACCCCATCATTCCTCTTTTGTGTGTGTTTCTTTCATGAC 371

Qy 964 GAGATGAGTGAATCAAGTTGGTTCGAATGTTTGTGGCCATCTCGGGATTTGTTTCTTAT 1023
```

```
Db 370 AAGATGATGGAGTAAGATTATAGCTATGCTGATGCCAATTTGGGATTTATGTCTAT 311
Qy 1024 GGTATCAGCATTTATGTCAATGATAGAAA 1052
Db 310 GGCCACCAATTATATGTTGATGCAAGAA 282
```

## RESULT 5

US-10-424-599-102519  
; Sequence 102519, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 102519  
; LENGTH: 2214  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_63592C.1  
US-10-424-599-102519

Query Match 24.5%; Score 280.4; DB 16; Length 2214;

Best Local Similarity 57.0%; Pred. No. 2.9e-66;  
Matches 583; Conservative 0; Mismatches 416; Indels 23; Gaps 3;

```
Qy 70 TGGAGGTTAAGAGTGTCTCTATGTCACTCTCTCTTAGCTGGAGAGACAATAGCCACT 129
Db 408 TGGTGGTGTCTTGTGACACTCAGCATAGCCCTTCTTATAGTTGGCAATCTGCTGCTGT 467

Qy 130 CTCTTAGGTAGACTTTACTACGAAAAAGCGGTAAAAAGCACATGGCTCGAAACCTTGGTT 189
Db 468 ATCTTGGAAAGATTTTATTATGATCAGGGTGGAAATAGTAAATGGATGGCTACTCTAGTT 527

Qy 190 CAGTTGTAGGGTTTCTTTTAAACCTTCTTGTCTTATTTACTTTAAAGCTGAGCCGCTCC 249
Db 528 CAAACTGCTCGCTTCCCGAT--CTTGTCTCATTCATTTATTAACAATTCCTTCACCTCCAG 585

Qy 250 AAGACTAAAAACCATTAACAAAAAACTACTTCTTCTTGTGACACTATCTTTAGTGTAT 309
Db 586 AGGCTTCAACTTCTGCTTCACTCCCATCAAAATTAATCTTTGATATATTTGGTCT-- 643

Qy 310 ATTGGACTTGGCTTGTGTTGTGGACATTTGATTTTGTACTCATATTTGGGCTACTTTAC 369
Db 644 -----TTGGAGTCTTAAATTTGCTGTGACAAATATGATGTACTCCACTGGACTCTTATAC 696

Qy 370 CTTCCTGTCTCAACTTCTCTTTTGTATCTCTGCGTGGCAATTTGGCTTTTAAAGCGCTTTC 429
Db 697 CTCTCGGCTTCTACCTTATTCGCTGATTTGTGCAATCAGATTAGCTTTTAAATGCAAGTTTC 756

Qy 430 TCTTACTTCTTAACTCACAATAAAATCACAACCAATTTATCTCAATTCACATTTGTCTCTTA 489
Db 757 TCATATTTTATCAATCTCAAAAGTTTCACTGCTTGTATTAATACTCTACAGTGGTTCTC 816

Qy 490 ACCATATCTTCTACATTTCTTGTATCCAAATGAAACAGAAATCTCCCTCTTCTACTTTCA 549
Db 817 ACTTTATCTGTCGACTCTCTTGTCTTGTGTTTAAAGGAAACAGACAGATGAACCATCTGGTTCT-- 874

Qy 550 AAGTCCGACCGCAAGTCCAAAGTATGTGATTTGGATACATCTGCGGCTCGGTAGCTCAGCT 609
Db 875 -----CCAAAGGAAAGTACATTAATTTGGTTTCTTATGTACCTTGGAGCTTCTGCA 924

Qy 610 GGTATTCTCTGGTGTCTTCTTTTAAACAGATTAAGCGTTTCGAAAGATTTCTTAAAGAAATAC 669
```

Db 925 GTGACTCTCTTTTGGCTTCCCTCATGCAGCTGACCTTTGAGAAGGTTCTGAAGAAGGAA 984  
QY 670 ACATTCAAGGCTATTTTATAGACATGGCCACATATCCGCTATATGGTAGCTACTTGTGTAGTT 729  
Db 985 ACATTTTCTGTGTTTGGAAATGCAAACTACACATCATTTGCTTGGCTCTGGTCTCT 1044  
QY 730 GTGGTAGGACTTTTGGAAAGTGTGGGTGAAAGAGCTGAGTACAGAAATGGAAGAGTTT 789  
Db 1045 GTCATAGGCCCTATTTGCAAGTGGGAAATGGCGTACTTTTGCATGGAGAAATGGAGGGTTT 1104  
QY 790 CAACCTAGGAAAAAGCTCATACATTTTGTATAAATCATCGGTTCAACGATATCATGCGAGCT 849  
Db 1105 CAGAAAGGATATGTTGCTTATGATATGACTTTGGTTTGGACTTCAATAGCCTGGCAGGTA 1164  
QY 850 TGTATTGATGGAAGTGTGGTTTGTATATCGAAGTTTTCATCGCTTTTTCCTCAATGTCATA 909  
Db 1165 TGCTCTGTGGTGTGTTGGCTTGAATCTTCTAGTGTCTTCTCTACTCCAATGTTATA 1224  
QY 910 AGCACTCTTGTATTACAGTGTGGCTGTCTTGTGCTGTGTTGCTTCTTCCGTGATGAGATG 969  
Db 1225 AGCACAGTTTCTTTAGCCGTAACCTCCTATTGCTGTCTTATAGTTTTCATGATAAGATG 1284  
QY 970 AGTGAATCAAGTTGGTTGCAATGTTTGTGGCCATCTGGGATTTGTTCTTATGCTTAT 1029  
Db 1285 AATGGGGTGAAGATAATTTCTATGCTTTTGGCTCTATGGGTTTTCCTCTTATATTTAT 1344  
QY 1030 CAGCATTATGTCATATGATAGAAAGCCAGAAAGACCAAGAGCTTCTCAGTCTAAAGAA 1089  
Db 1345 CAGAAATATCTGATGATTCAGACACAGACATGCACAGCTGCTACTAAGTCCCAAT 1404  
QY 1090 GA 1091  
Db 1405 GA 1406

RESULT 6

US-10-416-898-9  
; Sequence 9, Application US/10416898  
; Publication No. US20040172670A1  
; GENERAL INFORMATION:  
; APPLICANT: Yale University  
; APPLICANT: Walker, Elsbeth  
; APPLICANT: Dellaporta, Stephen  
; TITLE OF INVENTION: MAIZE YELLOW STRIPE1 AND RELATED GENES  
; FILE REFERENCE: 44574-5106-US  
; CURRENT APPLICATION NUMBER: US/10/416,898  
; CURRENT FILING DATE: 2003-05-16  
; PRIOR APPLICATION NUMBER: PCT/US01/43101  
; PRIOR FILING DATE: 2001-11-16  
; PRIOR APPLICATION NUMBER: US 60/249,222  
; PRIOR FILING DATE: 2000-11-16  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 9  
; LENGTH: 83698  
; TYPE: DNA  
; ORGANISM: Zea mays  
US-10-416-898-9

Query Match

Best Local Similarity 54.4%; Pred. No. 9.1e-57;  
Matches 547; Conservative 0; Mismatches 438; Indels 21; Gaps 2;  
QY 70 TGGAGGTTAAGAGTGTCTCTATGTCACTCTCTCTAGCTGGAGAGACAATAGCCACT 129  
Db 67228 TGGTGAATCTGTTTTCATAGCACTCTCTCTCATCTCTGCTCAAGCCATTTCTGTT 67287  
QY 130 CTCCTTAGGTAGACTTTTACGAAAAAGGGGTAAAGCACATGGCTCGAAACCTTTGGTT 189  
Db 67288 CTCCTTGGTCGGTTTTTATTACAAATGAAGGTGGAACAGATAATGGATCTCTACTCTTGT 67347  
QY 190 CAGCTTGTAGGGTTTCCCTTTAAACCCCTTCTGCTATATATCTTAAAGCCTGAGCGTCC 249

Db 67348 CAAACTGGTGGCTTTCCCAATCTTTATCTCCCTCTTTCTCTCTCTCTCTCTCTCTCTCT 67407  
QY 250 AAGACTAAAAACCATTAACCAAAAAAACTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 309  
Db 67408 TCTTCT 67458  
QY 310 ATGGACTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 369  
Db 67459 CTTTCT 67518  
QY 370 CTTCT 429  
Db 67519 CTTCT 67578  
QY 430 TCTTACTTCT 489  
Db 67579 TATTATTACATCAATCTCTCAGAAATACCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 67638  
QY 490 ACCATATCTTCT 549  
Db 67639 TCTATCT 67696  
QY 550 AAGTCGGCAGCCCAAGTCCCAAGTATGTGATGATGATGATGATGATGATGATGATGATGAT 609  
Db 67697 CTTCT 67746  
QY 610 GGTATTCT 669  
Db 67747 ATCTATTCT 67806  
QY 670 ACATTCAAGGCTATTTTATGACATGCCCACATATCCGTCTATGCTAGCTACTTGTGTAGTT 729  
Db 67807 ACTCTCTCTATGGTTCTCGAGATGCAAAATCTATACGTCTTGTGGCTTCTTGTGTAGCG 67866  
QY 730 GTGTAGGACTTTTGGAAAGTGTGGGTGGAAGAAAGCTGAGTACAGAAATGGAAGAGTTT 789  
Db 67867 GTTATCGGATGTTTCGCAAGCGGGAATGGATGTTTGTGAGTGTGGAGATGGAAGAGTTT 67926  
QY 790 CAACTAGGAAAAAGCTCATACATTTTGTATAAATCAACATCGGTTCAACGATATCATGGCAAGCT 849  
Db 67927 CAGGAAGGTCAAGTCATTTATGTTTGTGACTTGGTTCGGGCGAGCGGTTCTGTCTCAATG 67986  
QY 850 TGTATTGATGGAAGTGTGGTTTGTATTATCGAAGTTTTCATCGCTTTTTCCTCAATGTCATA 909  
Db 67987 GTTGTGTAGGAGCGGTGTCTCTTATATTTCTGGTGTCTCTCTCTCTCTCTCTCTCTCT 68046  
QY 910 AGCACTCTTGTTTTACCAAGTGTGCTGTTCTTGTCTGTGTTCTCTCTCTCTCTCTCTCTCT 969  
Db 68047 AGTACGCTCTCACTCATTTGTTACGCTCTCTCGCGGCAATTTGCGGTGTTCATGACAAGCTG 68106  
QY 970 AGTGAATCAAGTGTGGTTCATGTTTGTGGCCATCTGGGATTTTGTCTTATGCTTAT 1029  
Db 68107 ACTGAGTTAAGATGTTGGGATGCCATCGCTCTTACAGGATTCACGTTTATATCTCTAC 68166  
QY 1030 CAGCATTATGTCATGATAGAAAGCCAGAAAGACCAAGAGCTTC 1075  
Db 68167 CAGAACTATCTTGATGACTTGAAGAGTACAAAGAGCAGCAGAGAGCTC 68212

RESULT 7

US-10-424-599-75161  
; Sequence 75161, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 285684

; SEQ ID NO 75161

; LENGTH: 997

; TYPE: DNA

; ORGANISM: Glycine max

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_38884C.1

US-10-424-599-75161

Query Match 19.0%; Score 217; DB 16; Length 997;

Best Local Similarity 56.0%; Pred. No. 1.7e-50;

Matches 490; Conservative 0; Mismatches 355; Indels 30; Gaps 3;

```
QY 61 AAGTACTCTTGGAGTTAAGAGTGTCTCTATGTCACTCTCTCTAGCTGGAGACA 120
DB 125 AAATACAAATGGTGGCTACGTGTCCCTCTACATAATCTTCTTGTAGGCCAATCT 184
QY 121 ATAGCCACTCTCTTAGGTAGACTTTACTACGAAAGGCGGTAAAGACATGCTCGAA 180
DB 185 CGAGCCACTCTTTGGGAGAGATTACTATGACAAATGGTGTATAGCAATGATGGCA 244
QY 181 ACCTTGGTTCAGCTTGTAGGTTTCTTTAAACCTTCTCTGCTATTATTACTT---AAAG 237
DB 245 ACAATTGTTCAATCAGCTGGATTCCCTGTACTTCTTCCACTCTTTTACTTCCCAAGA 304
QY 238 CTTGAGCCGCTCAAGACTAAACCACTTACCAAAAAAACTACTTCTTCC-----285
DB 305 CAAACACATGCAAGTTTAAACAAATCCCGCAATATGATTAATCTACAAACAAAA 364
QY 286 -----TTCTTGACACTATCTTTAGTGATATTTGGACTTGGCTTGTGTGCTGGACAT 339
DB 365 CCAAAATTTCTCCACCTTGTCTTCTCTATCTAGCTTTTGGCCTTAATTTCTCACAGGGGAC 424
QY 340 TGTATTTTGTACTCAATTTGGCTACTTTPACCTTCTGTCTCAACTTTCTTTGATCTCT 399
DB 425 AACTTGATGATTTCTATGGACTTTTATATCTCCTCTTCTTCCACTCTTCTACTATGT 484
QY 400 GCGTGGCAATTTGGCTTTTAAAGCGGTCTTCTTACTTCTTAACTCACAATAATCACA 459
DB 485 GCCACCAATATAGTCTTCAACGCGGTGTCTTCTTCTTCTCAATTTCCAGAAATTCACA 544
QY 460 CCATTATACTCAATTTCACTTGTCTCTTAACCATATCTTCTACACTTCTTGTATCCAA 519
DB 545 GCATTCAATCAATTTCTGTAGTCTCTTACCATAATCAGCTTCTGTCTTGCATCAAC 604
QY 520 CATGAACAGAAATCTCCTCTTCTACTTCAAGTCCGAGCCAAAGTCAAGTATGTGATT 579
DB 605 TCCGATTCCGACGAGGAC-----TCAACCGGTCTTTCCAGAGAAAGCATGTAAAT 655
QY 580 GGATACATCTGCGGGTGGTAGCTCAGCTGGTTATTTCTGTGGTCTTCTTTAACAGAT 639
DB 656 GGATTTCTTGCAACCAATGGTGCAATCAGCTACATTTCTTTGTACCTCTCTTTGTGCA 715
QY 640 TACGGTTTCGAAAGATTCTAAAGAAATPACACATTTCAAGGCTATTTTAGACATGGCCACA 699
DB 716 CTTTCTTTTCAAGATTATTAAGAGAAACCTTTCTGTCTGTGTGGACATGCAATTT 775
QY 700 TATCCGCTATTTGGTAGCTACTTGTGTAGTGTGGTAGGACTTTTGGAGTGGTGGTGG 759
DB 776 TACCACCTCTTCAATGCTACATGTCTGTGTAGTAGGGTTGTTTGCAGTGGAGAGTGG 835
QY 760 AAAAGCTGAGTACAGAAATGGAAGTTTCAACTAGGGAAGCTCATACATTTTGTGATA 819
DB 836 AAAAGTTTGAACAATAGATGAAGGATATGATAAGGGAAGTGTGGCTTTGTAATGACT 895
QY 820 AACATCGGTTTCAACGATATCATGGCAAGCTTGTGTTGATTGGAAGTGTGGTTTGAATATC 879
DB 896 GTACTATAGATTGCTGTGACATGGTATATATGCTCACTAGGATGCTAGGGGGGACGTTG 955
QY 880 GAAGTTTCATCGCTTTTTCAAATGTCATNAGCAC 914
DB 956 GAGGAGTCTGATTTGTCGGATGTGAGAAAGTAC 990
```

RESULT 8

US-10-437-963-43860/c

; Sequence 43860, Application US/10437963

; Publication No. US20040123343A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; APPLICANT: Wu, Wei

; APPLICANT: Boukharov, Andrey A.

; APPLICANT: Barbazuk, Brad

; APPLICANT: Li, Ping

; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53221)B

; CURRENT APPLICATION NUMBER: US/10/437,963

; CURRENT FILING DATE: 2003-05-14

; NUMBER OF SEQ ID NOS: 204966

; SEQ ID NO 43860

; LENGTH: 2381

; TYPE: DNA

; ORGANISM: Oryza sativa

; FEATURE:

; NAME/KEY: unsure

; LOCATION: (1)..(2381)

; OTHER INFORMATION: unsure at all n locations

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_46977C.1

US-10-437-963-43860

Query Match 17.2%; Score 197.2; DB 17; Length 2381;

Best Local Similarity 51.4%; Pred. No. 1.2e-44;

Matches 528; Conservative 0; Mismatches 473; Indels 27; Gaps 2;

```
QY 70 TGGAGGTTAAGAGTGTCTCTATGTCACTCTCTCTAGCTGGAGAGACAATAGCCACT 129
DB 2246 TGGTGGCGATGTGGCAGTGGAGCTTCTTCTCTATCGCCGTGACACATCGGCGACA 2187
QY 130 CTCTTAGGTAGACTTTACTACGAAAAAGCGGTAAAAAGACATGGCTCGAAAACTTGTT 189
DB 2186 CTGCTGGCGAGATACTACTACCCAAAGCGCGCGAGCAAGTGGATATCGGCGTTCGTG 2127
QY 190 CAGCTTGTAGGGTTCTTTAAACCTTCTCTGTCTATTATTACTTAAAGCTGAGCCCTCC 249
DB 2126 CGGACGCGCGGCTTCCCGATACTGTTCTTCAACCTCTCTTCTTCCCTCGAAGTCACCG 2067
QY 250 AAGACTAAACCACTTACCAAAAAAACTACTTCTTCTTCTTGACACTATCTTTAGTGTAT 309
DB 2066 TCCTCTGCACCAACA-----CTCCCATGGCCAGCTCGCGGTATATAC 2022
QY 310 ATTGGACTTGGCTTGTGTGTGGACATTTGATTTTGTACTCAITTTGGGCTACTTTTAC 369
DB 2021 ATGCTCTCGGCTCATCATCGCGCGAGACATGATGATGATGATGATGATGATGATGATG 1962
QY 370 CTTCTGTCTCAACTTCTTCTTTGATCTCTGCGTCGCAATTTGCTTTTAAAGCGCTCTTC 429
DB 1961 CTCCCGCTTCGACCTACTCGCTCACTGTGCGCAGCCAGCTCGCGTTCAATGTCTCTTC 1902
QY 430 TCTTACTTCTTAACTCACAATAAAATCACACCACTTTATATCTCAATTCACATTTGTTCTCTTA 489
DB 1901 TCATACGTGTCTCACTCCGAGAGGTCACCTCTCTGATATTCACTCCGTCGTCTGCTC 1842
QY 490 ACCATATCTTCTACACTTCTTGTATTCCAAACATGAACCAAAATCTCCCTCTCTACTTCA 549
DB 1841 ACCATGTCCGTTCTCGCTCATCGGAGTCAGCAAGGA-----GTCTCAGGGGGTCT 1794
QY 550 AAGTCCGACGCCAAGTCCAAAGTATGATTTGGATACATCTGCGGGTTCGCTAGCTCAGCT 609
DB 1793 ACCGCGCTCTCGGAGGGAAGTATCTGCTCGTTCGTTCTGCTGAGCTGGGGGGCGCTGTCG 1734
QY 610 GGTATTCTCTGGTGTCTTTCTTTAAACAGATTACGCGTTTCGAAAGAGATTCTTAAAGAAATAC 669
```

Db 1733 ACTTACTCGTGTGATCTCGCGCTGATCGAGCTCACCTTCGAGACCATCATCAAGAAGCAC 1674  
Qy 670 ACATTCAAGAGCTATTTTAGACATGGCCACATATCCGCTCTATGTGTAGTACTCTGTGTAGTT 729  
Db 1673 ACCTTCTCAGCCGCTCTCAACATGCGAGATCTACACGGCGCTCGTGGCGAGCGCGGTG 1614  
Qy 730 GTGGTAGGACTTTTGGAAAGTGTGGGTGGAAGAAAGCTGAGTACAGAAATGGAAGAGTTT 789  
Db 1613 GTGGTCGGGCTGTTCGCGAGCGGAGTGGAGTCTGAGGGGGGAGATGAACGCGTTC 1554  
Qy 790 CAACTAGGGAAGAGCTCATACATTTTGATAAAACATCGTTCAACGATATCATGGCAAGCT 849  
Db 1553 AGTTCGGGGCAGTTCCTACCTGATGACGCTCTGTGGCGCGCGTGTGTCGGCAGGTG 1494  
Qy 850 TGTTCGATTGGAAGTGTGTGTTGATTATCGAAGTTTCATCGCTTTTTCCTCAATGTCATA 909  
Db 1493 GCCAACATCGGGGTGTCGGCTCACTTCGAGGTGTGGCGCTCTCTCCACAGTGATC 1434  
Qy 910 AGCACTCTTTGTTTACAGTTGCGCTGTTCCTGCTGTGTGTTCTTCCTCGTGATGAGATG 969  
Db 1433 AGCACGGTGTCACTCGCGGTGTCATCCCGTTCCTCGCGTGTGTGTTCCACGACAGATG 1374  
Qy 970 AGTGGATCAAGTTGGTTCGAATGTTTTCGCGCATCTGGGATTTGTTCTTATGTTAT 1029  
Db 1373 AACGGGGTGAAGATCGTGGCCATGCTGATTCGAATTTGGGGATTTATTTTCGTATCTGTTT 1314  
Qy 1030 CAGCATTTATGTCATATGATAAGAGCCAGAGAACCAAGAGCTTCCTCAGTCTAAGAA 1089  
Db 1313 CAGCATTTATGATGCGAAGAAAGCGAAGAGGCTTCATCGGGTGATAGTTCGAGGC 1254  
Qy 1090 GAAGAAGA 1097  
Db 1253 CAAAGAAGA 1246

## RESULT 9

US-10-767-701-9000  
; Sequence 9000, Application US/10767701  
; Publication No. US20040172684A1  
; GENERAL INFORMATION:  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement  
; FILE REFERENCE: 38-21(53535)B  
; CURRENT APPLICATION NUMBER: US/10/767,701  
; CURRENT FILING DATE: 2004-01-29  
; NUMBER OF SEQ ID NOS: 63128  
; SEQ ID NO 9000  
; LENGTH: 728  
; TYPE: DNA  
; ORGANISM: Sorghum bicolor  
; FEATURE:  
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS69833\_1  
US-10-767-701-9000

Query Match 14.7%; Score 168.6; DB 17; Length 728;  
Best Local Similarity 60.7%; Pred. No. 7.3e-37;  
Matches 276; Conservative 0; Mismatches 179; Indels 0; Gaps 0;

Qy 598 GGATGCTCAGCTGGTTATTTCTCTGGTGTCTTTTAAACAGATTACGCGTTTCGAAAAGATT 657  
Db 1 GGAGCTCAGCCNACTACTCACTAATCTCTCTAATCGAAGTCACATTTGAGAGGTT 60  
Qy 658 CTAAGAATAACACATTCAGGCTATTTTAGACATGGCCACATATCCGCTATGTAGTCT 717  
Db 61 ATTAAGAAGAAACGTTCTCAGTCGTTGTAATATGCATATATACACACTAGTGGCA 120  
Qy 718 ACTTGTGTAGTTGTGTAGACTTTTGGAAAGTGTGGTGGAAAGCTGATGACAGAA 777  
Db 121 ACAATAGCTTCTCTTGTGGGTTATTTGCAAGCGGTGAGTGGAAAGACTTTAGAGGAGAG 180

Qy 778 ATGGAAGAGTTTCAACTAGGAAAAGCTCATACATTTTGTATAAACAATCGGTTCAACGATA 837  
Db 181 ATGCATGCCCTTCAGCTCAGGAGGTTGTCTATGTGATGACACTTCTATGGAGCTGCTGA 240  
Qy 838 TCATGGCAAGCTGTTTGTGATGGAAGTGTGGTTGTGATTTATCGAAGTTTTCATCGCTTTT 897  
Db 241 TCTTGGCAGATAGCTTCTGTAGGAGTGGTGGGTTTGTATCTTTTGTGTGTCATCACTCTTT 300  
Qy 898 TCCAATGTCATAAGCACTCTTTTGTATACCAGTTGTGCGTGTCTTGTCTGTCTCTTC 957  
Db 301 TCAATGTGATTAAGCACACTAGCTCTACCCATCATTTCCAGTTTTTGTCTGTGATTTTCTTC 360  
Qy 958 CGTGATGAGATGAGTGGAAATCAAGTTGGTTGCAATGTTTTCGCCCATCTGGGGAATTTGT 1017  
Db 361 CACGACAAGATGAATGGAGTGAAGATTATGCTATGTTGATGCGCATCTGGGGAATTTGT 420  
Qy 1018 TCTTATGGTTATCAGCATTTATGTCATGATAGATAAA 1052  
Db 421 TCATATGGATACCAATTATATATTAGTGACAAGAA 455

## RESULT 10

US-10-767-701-6806  
; Sequence 6806, Application US/10767701  
; Publication No. US20040172684A1  
; GENERAL INFORMATION:  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement  
; FILE REFERENCE: 38-21(53535)B  
; CURRENT APPLICATION NUMBER: US/10/767,701  
; CURRENT FILING DATE: 2004-01-29  
; NUMBER OF SEQ ID NOS: 63128  
; SEQ ID NO 6806  
; LENGTH: 747  
; TYPE: DNA  
; ORGANISM: Sorghum bicolor  
; FEATURE:  
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS55663\_1  
US-10-767-701-6806

Query Match 14.2%; Score 163; DB 17; Length 747;  
Best Local Similarity 60.5%; Pred. No. 2.9e-35;  
Matches 268; Conservative 0; Mismatches 175; Indels 0; Gaps 0;

Qy 69 TTGAGGTTAAGAGTGTCTCTATGTGTCACCTCTCTCTTAGCTGGAGAGACAATAGCCAC 128  
Db 304 TTGGTGGTTGATGGTAGCTCTGAACATTTTCTTCTCTGTTGCTGTCAGACAGCATCGAC 363  
Qy 129 TCTCTTAGGTAGACTTTTACTAGAAAAGGCGGTAAAGACATCGGTCGAAAACCTTGGT 188  
Db 364 ACTCCTTGGTAGGTTCTACTACCAAGGTGGAAACAGCAATGGATGTCCACATTTGT 423  
Qy 189 TCAGCTTGTAGGTTTCTCTTAAACCTTCTTGTATTTACTTAAAGCTGAGCCGTC 248  
Db 424 CCAAAACCGCTGGCTTCCAGTATTTGTTGTCGCCCTATATCTGTCCGTTCAAAGTCGCC 483  
Qy 249 CAAGACTAAAACCAATTACCAAAAAAATCTCTTCTCTTTTGACACTATCTTTAGTGTA 308  
Db 484 TTCTACTCAAAACCAACCAAGTAAACCTGAGACTTCTGTCCACCAAGATTTACTTTATA 543  
Qy 309 TATTGGACTTGGCTTGTGTTGCTGGACATTTGTTACTATTTTGTACTCATTTGGGCTACTTTA 368  
Db 544 CATTTGGCTTGGGCTCATCATTTGCTGCGATGACTTGTATTTCTATGCTATGGCTCTCTGTA 603  
Qy 369 CTTTCTCTCTCAACTTTCTCTTTGATCTCTGCGTCGCAATTTGGCTTTTAAAGCCGCTCT 428  
Db 604 CTTTCTCTGTACCAATATTCGCTCATATGCGCTAGTCAGCTGGCCCTTCAATGCTGCTCT 663  
Qy 429 CTCCTTACTTCTTAAACTCAGAAAAATCACACCAATTTATCTCAATTCATCTTGTCTCTT 488

Db 664 CTCTTATGTCCTGAATGCTCAAAAGTTCACTGCAATTGATCTCAACTCGGTAACTCTCT 723  
Qy 489 AACCATATCTTACACTTCTTG 511  
Db 724 TACCTTTTCTGTCGCACTTCTTG 746

## RESULT 11

US-09-770-152-44  
; Sequence 44, Application US/09770152  
; Publication No. US20020040489A1  
; GENERAL INFORMATION:  
; APPLICANT: Goriach, Jörn  
; APPLICANT: An, Yong-Qiang  
; APPLICANT: Hamilton, Carol M.  
; APPLICANT: Price, Jennifer L.  
; APPLICANT: Raines, Tracy M.  
; APPLICANT: Yu, Yang  
; APPLICANT: Rameaka, Joshua G.  
; APPLICANT: Page, Amy  
; APPLICANT: Matthew, Abraham V.  
; APPLICANT: Ledford, Brooke L.  
; APPLICANT: Woessner, Jeffrey P.  
; APPLICANT: Haas, William David  
; APPLICANT: Garcia, Carlos A.  
; APPLICANT: Kricke, Maja  
; APPLICANT: Slader, Ted  
; APPLICANT: Davis, Keith R.  
; APPLICANT: Allen, Keith  
; APPLICANT: Hoffman, Neil  
; APPLICANT: Hurban, Patrick  
; TITLE OF INVENTION: Expressed Sequences of Arabidopsis  
; TITLE OF INVENTION: Thalana  
; FILE REFERENCE: 2025US (PARA-014PRV)  
; CURRENT APPLICATION NUMBER: US/09/770,152  
; CURRENT FILING DATE: 2001-01-26  
; PRIOR APPLICATION NUMBER: 60/178,503  
; PRIOR FILING DATE: 2000-01-27  
; NUMBER OF SEQ ID NOS: 999  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 44  
; LENGTH: 592  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-09-770-152-44

Query Match 13.5%; Score 154.6; DB 9; Length 592;  
Best Local Similarity 56.9%; Pred. No. 6e-33;  
Matches 283; Conservative 0; Mismatches 214; Indels 0; Gaps 0;  
Qy 28 GAAGGGAATTTTCAGAGAGAGAGAGTCAAGTACTCTTGAGGTTAAGAGTGTCT 87  
Db 94 GAAAGTGAATCATTTTTCAGTACCTCAACGAGAACTGTAAAGAGTGGCTCGGTGTCTCC 153  
Qy 88 CTCTATGTCACTCTCTAGCTGGAGACAATAGCACTCTCTTAGGTAGACTTTAC 147  
Db 154 ATATACGCAATCTTGTCTATCTTGTGCCAACCACTTGCTACAGTTCTGGGTAGACTGTAC 213  
Qy 148 TACGAAAGGGGTAAAGCACATGGCTCGAAACCTTTGGTTCAGCTTTGTAGGGTTTCCT 207  
Db 214 TATGAAATGAGGGAAGAACACATATGTTGTAACACTTCTTCAACTCATTTGGCTTCCT 273  
Qy 208 TTAACCTTCTTGTATATTACTTAAAGCTGAGCGGTCAAGACTAAACATATACC 267  
Db 274 GTACTGATTTCTTCCGGTCTTTTCTCGAATCAGGCAACCAATCAACAGATACAAAT 333  
Qy 268 AAAAAAATCTCTTCTCTTGCACATATCTTTAGTGTATATTGGACTTTGGCTTGTCT 327  
Db 334 TTCAGTCACTCCCTTCTCTTACACACCTTCGATCGGTTTACTTGTGCACTGGACTGCTA 393  
Qy 328 GTTCTGGACATGTATTTTGTACTCAFTTGGGCTACTTTACTCTTCTGCTCAACTTTC 387

Db 394 GTGTCTGTTATGCTTATTTGTTCTGCAAGTTGGGTGCTTTACTTACCAGTCTCTACTTTC 453  
Qy 388 TCTTTGATCTCTCGTCGCAATTGGCTTTTAAAGCGCTCTCTCTTACTTACTTAAACTCA 447  
Db 454 TCCCTCATCTTGGCCTCACAGTTGGCCTTCACTGCCTTTTCTCATATTTCCCTTAACTCG 513  
Qy 448 CAAAAAATCACACCATTTATATCAATTCATCTTCTTCTTAAACATATCTTCTACACTT 507  
Db 514 CAAAAGTTCACTCTTTGATAGTCAATTTCTTTGTTTCTCTCTAGGGTTTCTCTGCGCTC 573  
Qy 508 CTTGTTATCCAACATGA 524  
Db 574 CTCGTGGTCAACACTGA 590

## RESULT 12

US-10-437-963-83900  
; Sequence 83900, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 83900  
; LENGTH: 2605  
; TYPE: DNA  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_83188C.1  
US-10-437-963-83900

Query Match 11.4%; Score 130.6; DB 17; Length 2605;  
Best Local Similarity 49.7%; Pred. No. 9.2e-26;  
Matches 368; Conservative 0; Mismatches 364; Indels 9; Gaps 1;  
Qy 305 TGTATATTGCACTTGGCTTGTCTGTCTGGACATTTGTATTTGTACTCATTTGGGCTAC 364  
Db 1526 TCTACGTGGCCTCGGATCATCATCGGCTTCGACACACCTGATGTACTCGTACGCGTGC 1585  
Qy 365 TTTACCTTCTGTCTCACTTCTCTTTGATCTCTGGTCGCAATTCGGCTTTTAAACGCG 424  
Db 1586 AGTACCTGCGCGGTGCACCTTCTCGCTCGTGGCGGACGACGCTCGGCTTCAACTCCG 1645  
Qy 425 TCTTCTTACTTCTTAACTCAAAAAATCAACCAATTTATCTTACTCAATTTCACTTGTTC 484  
Db 1646 TCACCTCAGCTCATCAACGCGAGCGGTTACGGTGTGTATGTCGCCAATCTCCGTCTGTCG 1705  
Qy 485 TCTTAAACCATATCTTCTACACTTCTTGTATTCACAAATGAACCAAGATCTCCCTTCTTA 544  
Db 1706 TGCTCACCTTCTCGGCGCGCTGCTCGGCATC-----GGGGCCTCTCTCCGACGAGA 1756  
Qy 545 CTTCAAGTCGAGCGCAAGTCCAAAGTATGTGATTGGATACATCTGCGCGTTCGGTAGCT 604  
Db 1757 CCGCAGACGCGTCCGAGGGGCAAGTACCCCGCGGGTTCGCCCTGACGCTCGCCGCT 1816  
Qy 605 CAGCTGTTATTTCTCTGGTCTTCTTTTAAACAGATTACGGGTTGCAAAAGATTCTAAAGA 664  
Db 1817 CCGCGTGTTCGCGCTCATCTGTGCTGTTTCGAGGCCACTTCGAGAAGTGTGTCGGA 1876  
Qy 665 AATACACATTCAGGCTATTTTATAGACATGCGCCATATTCGTCATATGGTAGCTACTTGTG 724  
Db 1877 CGCGGACGCTCCGCTGGGTTCTCGCGCGCAGCTGTGTGGACCAAGGTGTGCGCTCGACGG 1936



```
QY 725 TAGTTGTGTAGGACTTTTGGAAAGTGTGGTGGAAAAAGCTGAGTACAGAAATGGAAG 784
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1937 TGTCCGGCGTGGGCGTCTCGCGTCGGGGGACTGGAGGACGATACCGCGGAGATGGCGG 1996
QY 785 AGTTTCAACTAGGAAAAAGCTCATACATTTTGTATAAACATCGGTTCAAGCATATCATGGC 844
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1997 CGTTCAAGGACGGGAGGGGAGGATGACGTGGCGACGCTGTGTGGGACGGCGGTGTCTGGC 2056
QY 845 AAGCTTGTGTGATTGGAAGTGTGGTTTGATTATCGAAGTTTTCATCGCTTTTTCCAATG 904
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2057 AGTGATGGCGGTGGGCTCGCTCGGCTGATCGTAGGGTGTGTCGCTGTTCGCCAAGC 2116
QY 905 TCATAAGCACTCTTTGTACCAAGTTGTGCCTGTTCTTGCTGTGTGTCCTTTCCTCGGTGATG 964
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2117 TGACGGGCACGCTGTCTGCTGCCGTGTTCCGCGTGTTCGCCGTGGCGGTTCGGGGACA 2176
QY 965 AGATGAGTGAATCAAGTTGGTTGCAATGTTTTTGGCCATCTGGGATTTGTTCTTATG 1024
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2177 GGATGACCGGGATCAAGGCCGTATCCATGCTCATGGCCGTCTGGGGTTTCTCTGTAGC 2236
QY 1025 GTTATCAGCATATGTCAATG 1045
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2237 CGTACCAGCATACATCGAGC 2257
```

```
RESULT 13
US-10-424-599-9213
; Sequence 9213, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 9213
; LENGTH: 792
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(792)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_108329C.1
US-10-424-599-9213
```

```
Query Match 11.2%; Score 128.4; DB 16; Length 792;
Best Local Similarity 56.0%; Pred. No. 1.9e-25;
Matches 271; Conservative 0; Mismatches 201; Indels 12; Gaps 1;

QY 59 ACAAGTACTCTTGGAGTTAAGAGTCTCTCTATGTCACTCTCTTAGTGGAGAGA 118
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
108 AGAATATACATACGGTGTGCGTGTGTCATTTATACATCTCTTCTCGTTTGCCTAAT 167
QY 119 CAATAGCCACTCTCTTAGGTAGACTTTTACTAGCAAAAAGCGGTAAAAGCACATGGCTCG 178
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
168 GTTCAGCCACTCTTTTGGGAAGTTTACTTTTGACAAAGTGTGTAAGCAATGATG 227
QY 179 AAACCTTGGTTGAGTTGAGGGTTTCCTTTAAACCTTCCTTGTCTATTATTACT----- 232
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
228 AAGCATTTTCAATCAGCTGGATTCCCTATACCAATTCACCTCTATTTTACTGTCAA 287
QY 233 -----TAAAGCTTGAGCGTCCAAGACTAAACACCATTTACCAAAAACACTCTTCTCT 286
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
288 AACATGCAAAATCAACAAACGATCTCTAGTAATAATGATACCTCCGAAACCAAAAC 347
```

```
QY 287 TCTTGACACTATCTTTAGTGTATATTTGGACTTGGCTTGTCTTGTGTGACATTGTATTT 346
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
348 TTCTCATTATATTTCTCTGTACCTAGTTTGTGGATTATTGTTAGCAGCATAGACTTTA 407
QY 347 TGTACTCATTTGGGCTACTTTTACCTTCTGTCTCAACTTCTCTTTGATCTCTGGGTCCG 406
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
408 TGATGTCATGTGGACTTTTATACCTTCTCTTCCACTTTTTCACATTTATGTGATCCG 467
QY 407 AATTGGCTTTTAAACGCGCTTCTTCTTACTTCTAAACTCAACAAAAATCACACCATTTA 466
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
468 AATTGGGCTTCATTCAGATGTGACTTCTTCATCATCAATCCCAAAGTTCACTGCACTCA 527
QY 467 TACTCAATTCACTTGTCTTCTTAAACCATATCTTCTACACTTCTTGTATCCAAACATGAAC 526
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
528 TATTAACACTCTGTAGCTGTCTTAGCATGTGAGTTATCTTAATGCACTCAACCCCTGAAT 587
QY 527 CAGA 530
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
588 CTGA 591
```

```
RESULT 14
US-10-425-114-28473
; Sequence 28473, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Kovalic, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 28473
; LENGTH: 1779
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4757-034-H3_FLI
US-10-425-114-28473
```

```
Query Match 11.1%; Score 127; DB 16; Length 1779;
Best Local Similarity 50.1%; Pred. No. 7.7e-25;
Matches 350; Conservative 0; Mismatches 340; Indels 9; Gaps 1;

QY 347 TGTACTCATTTGGGCTACTTTTACCTTCTGTCCTGTCCTCAACTTCTCTTTGATCTCTGCGTCG 406
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
817 TGTACTCTGACGCTGACGTACCTTGGCGGTGTCACCTTCTCGCTCTGGCGCGGACGC 876
QY 407 AATTGGCTTTTAAACGCGCTTCTTCTTACTTCTTAAACTCAAAAAATCACACCATTTA 466
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
877 AGCTGGGCTTCAACGCCATCATCTCGGCTCATCAACGCGCAGCGTTTCAGCGCGCTGA 936
QY 467 TACTCAATTCACTTGTCTTCTTAAACCATATCTTCTACACTTCTTGTATCCAAACATGAAC 526
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
937 TCGCCAACTCCGTGGTGTGTCCTACCTTCTCCGCGCATCTCTCGGCGTC-----G 987
QY 527 CAGAATCTCCCTTCTTACTTTCAAAGTCCGAGCAAGTCCCAAGTATGTGATGGATACA 586
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
988 GATCTCTCCGACGAGACCTCCAGCAACGTCCGCGGCGCAAGTACCCGCTGGGTTCCG 1047
QY 587 TCTCGCGGTGCGTAGCTCAGCTGCTTATCTCTGGTGTCTTCTTAAAGATTACGCGT 646
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1048 TCCTCGTGTGCGCGCTCGCGGTGTTCGGGCTCATTTTGTCTCTTTCGAGCTCACCT 1107
QY 647 TCGAAAAGATTCTTAAAGAAATACACATTCAAAGGCTATTTTATAGACATGGCCACATTCGCT 706
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1108 TCGAAGAGGTGCTCAGGGTGCAGCGCGCGGTGGGTGCTCGGATGCAGATGCACACCA 1167
```

```
Qy 707 CTATGTTAGTACTTGTGTAGTCTTGTGTAGTACTTTTGGAGTGTGGTGGAAAAAGC 766
Db 1168 ACCTGGTGGCGTGGCGGTGTCCGTGGGGGGCTGCTCGCTCGGGGGACTGGCGGACGA 1227
Qy 767 TGAGTACAGAAATGGAAGATTTCAACTAGGGAAGCTCATACATTTTGATAAACATCG 826
Db 1228 TCCGGGGAGATGGCGTCGTTCAAGACGGGAGGACGAGGTACGTGCTGACGCTGGTGG 1287
Qy 827 GTTCAACGATATCATGGCAAGCTTGTGTGATGGAAGTGTGGTGTGATATCGAAAGTTT 886
Db 1288 GCACGGCGGTGTGGTGGCAGCAGCGCGGTGGCTTGGTGGCGCTGATCATGAGGGTGT 1347
Qy 887 CATCGCTTTTCCAAATGTCATAGCACTCTTTGTGTACCAAGTGTGCGCTTCTTGCTG 946
Db 1348 CGTGGCTTTCCGCAAGCTGACGTGCACGCTGGCGCTGCGCTGGTGCAGTGTTCGG 1407
Qy 947 TTGTCTTCTCGGTGATGAGTGAATCAAGTTGGTTGCAATGTTTTTGGCCATCT 1006
Db 1408 TGGCGCTGTTGGGGACAGGATGACCGGTATAAGATCGTGGCCATGCTCATGGCGGTAC 1467
Qy 1007 GGGGATTTGTTCTTATGTTATCAGCATTTATGCAATG 1045
Db 1468 GGGGGTTCCTCTCTCTACATGTACCAGCACTACATCGACG 1506

RESULT 15
US-10-425-114-12485
; Sequence 12485, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 12485
; LENGTH: 1386
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 701179255_FLI
US-10-425-114-12485

Query Match 11.0%; Score 125.4; DB 16; Length 1386;
Best Local Similarity 49.9%; Pred. No. 1.9e-24;
Matches 349; Conservative 0; Mismatches 341; Indels 9; Gaps 1;

Qy 347 TGTACTCATTTGGGCTACTTTTACCTTCCTGCTCAACTTCTCTTTGATCTCTGCTGC 406
Db 453 TGTACTCATGCGCTGAGTACCTTCCCGGTGTCACCTTCTCGCTCGGCGGACGC 512
Qy 407 AATTGGCTTTTAAACGCCCTCTTCTTACTTCTTAACTCAAAAATCACACCATTTA 466
Db 513 AGCTGGGCTTCACGCCATCACCCTCGCGCTCATCAACGCGAGCGGTTCAAGGCTGA 572
Qy 467 TACTCAATTCAGTTGTTCTTTAACCATATCTTTACACTTCTTGTATCCAAATGAAC 526
Db 573 TCGCAACTCCGCTGCTGCTGCTCACTTCTCGCGCGATCTCGGCGTC-----G 623
Qy 527 CAGATCTCCCTTCTTACTTCAAGTCCGAGCCCAAGTCCCAAGTATGATGATGATACA 586
Db 624 GATCTCTCTCGACGAGACCTTCACCAACGTCGCGCGGCAAGTACCACCTGGGGTTCG 683
Qy 587 TCTGCGCGGTGCGGTAGCTCAGCTGTTATTTCTCTGGTGTCTTTCTTTAAACAGATTACGCGT 646
```

```
Db 684 TACTCGTGTGGCGCCTCGCGGTGTTCGCGCTCACTTTTGTGCCCTCTTCAGCTCACCT 743
Qy 647 TCGAAAGATTTAAAGAAATACACATTCAAGGCTATTTTAGACATGGCCACATATCCGT 706
Db 744 TCGAAGAGTGGTCAAGGTCGACGCGCGGTGGTGGCTGCGGATGCAGATGCACACCA 803
Qy 707 CTATGGTAGCTACTTGTGTAGTTGTGTAGGACTTTTGTGAAAGTGGTGGGTGAAAAAGC 766
Db 804 ACCTGGTGGCGTGGCGGTGTCGTGGCGGGCTGCTCGCTCGGGGACTGGCGGACGA 863
Qy 767 TGAGTACAGAAATGGAAGAGTTTCAACTAGGGAAGCTCATACATTTTGAATAAACATCG 826
Db 864 TCCGCGGGAGATGGCGTCGTTCAAAGACGCGAGGACGAGGTACGTGCTGACGCTGGTGG 923
Qy 827 GTTCAACGATATCATGGCAAGCTTGTGTGATTTGGAAGTGTGGTGTGATTTATCGAAGTTT 886
Db 924 GCACGGCGGTGTCGTGGCAGCAGCGCGCTGGGCTTGGTGGCGCTGATCATGAGGGTGT 983
Qy 887 CATCGCTTTTTTCCAATGTCTAAGCACTCTTTTGTTTTACCAGTTGTGCCCTGTTCTTGCTG 946
Db 984 CGTGGCTGTTGGGAACTGACGTGCACGCTGGCGCTGCCGCTGGTGGCCAGTGTTTGCGG 1043
Qy 947 TTGTCTTCTTCCGCTGATGAGATGAGTGAATCAAGTTGGTGGTGAATGTTTGGCCATCT 1006
Db 1044 TGGCGCTGTTCCGGGACAGGATGACCGGTATAAAGATCGTGGCCATGCTCATGGCGGTAT 1103
Qy 1007 GGGGATTTGTTCTTATGTTTATCAGCATTTATGTCATG 1045
Db 1104 GGGGGTTCCTCTCTCTACATGTACCAGCACTACATCGACG 1142
```

Search completed: November 1, 2004, 23:39:57  
Job time : 599.039 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 1, 2004, 05:48:03 ; Search time 105.63 Seconds  
(without alignments)  
7704.726 Million cell updates/sec

Title: US-09-913-767-3

Perfect score: 1145

Sequence: 1 cgtgttcattgtgtgtgtg.....caagtttaggaagatcca 1145

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.\*

1: /cgn2\_6/prodata/1/ina/5A COMB.seq.\*

2: /cgn2\_6/prodata/1/ina/5B COMB.seq.\*

3: /cgn2\_6/prodata/1/ina/6A COMB.seq.\*

4: /cgn2\_6/prodata/1/ina/6B COMB.seq.\*

5: /cgn2\_6/prodata/1/ina/PCBUS COMB.seq.\*

6: /cgn2\_6/prodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	74.4	6.5	7218	1	US-08-232-463-14
2	45.6	4.0	399	4	Sequence 8976, Ap
3	45	3.9	3090	4	Sequence 5, Appl
4	42.6	3.7	3081	4	Sequence 9, Appl
5	41.6	3.6	1141	4	Sequence 22, Appl
6	41	3.6	1141	4	Sequence 22, Appl
7	40.2	3.5	500	3	Sequence 2, Appl
8	39.2	3.4	832	4	Sequence 2813, Ap
9	38.2	3.3	832	4	Sequence 2813, Ap
10	38.2	3.3	5935	3	Sequence 17, Appl
11	38.2	3.3	5935	3	Sequence 29, Appl
12	38.2	3.3	5935	3	Sequence 29, Appl
13	36.8	3.2	843	4	Sequence 5526, Ap
14	36.8	3.2	5092	4	Sequence 153, App
15	36.6	3.2	19233	4	Sequence 45, Appl
16	36.2	3.2	582	4	Sequence 3, Appl
17	35.8	3.1	1608	4	Sequence 1772, Ap
18	35.8	3.1	14759	4	Sequence 1, Appl
19	35.6	3.1	3117	4	Sequence 275, App
20	35.4	3.1	1447	4	Sequence 27, Appl
21	35.2	3.1	289	3	Sequence 17, Appl
22	35.2	3.1	289	3	Sequence 17, Appl
23	34.8	3.0	10619	4	Sequence 4, Appl
24	34.6	3.0	10007	3	Sequence 13, Appl
25	34.4	3.0	573	3	Sequence 440, App
26	34.4	3.0	789	4	Sequence 1652, Ap
27	34.4	3.0	879	4	Sequence 87, Appl

28	34.4	3.0	933	4	US-09-107-532A-78	Sequence 78, Appl
29	34.4	3.0	18073	3	US-09-078-294-12	Sequence 12, Appl
30	34.2	3.0	455	4	US-09-232-785-296	Sequence 296, App
31	34.2	3.0	722	4	US-09-270-767-8029	Sequence 8029, Ap
32	34.2	3.0	722	4	US-09-270-767-23311	Sequence 23311, A
33	34.2	3.0	1395	4	US-09-134-000C-2349	Sequence 2349, Ap
34	34.2	3.0	2834	1	US-08-276-151-8	Sequence 8, Appl
35	34.2	3.0	3054	4	US-09-762-724-11	Sequence 11, Appl
36	34.2	3.0	3073	1	US-07-688-352C-31	Sequence 31, Appl
37	34.2	3.0	3073	2	US-08-474-379C-31	Sequence 31, Appl
38	34.2	3.0	3073	3	US-09-146-249A-31	Sequence 31, Appl
39	34.2	3.0	3073	3	US-08-206-188B-31	Sequence 31, Appl
40	34.2	3.0	3073	5	PCT-US91-02714-30	Sequence 30, Appl
41	34.2	3.0	3084	4	US-09-763-724-7	Sequence 7, Appl
42	34.2	3.0	8093	4	US-10-204-708-32	Sequence 32, Appl
43	34	3.0	747	4	US-09-134-000C-2966	Sequence 2966, Ap
44	34	3.0	1170	4	US-09-248-796A-9079	Sequence 9079, Ap
45	34	3.0	1298	3	US-08-948-703-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1  
US-08-232-463-14  
; Sequence 14, Application US/08232463  
; Patent No. 5670367  
; GENERAL INFORMATION:  
; APPLICANT: DORNER, F.  
; APPLICANT: SCHREIFLINGER, F.  
; APPLICANT: FALKNER, F. G.  
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 1800 Diagonal Road, Suite 500  
; CITY: Alexandria  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22313-0299  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/232,463  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/935,313  
; FILING DATE:  
; APPLICATION NUMBER: EP 91 114 300.6  
; FILING DATE: 26-AUG-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BENT, Stephen A.  
; REGISTRATION NUMBER: 29,768  
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703)836-9300  
; TELEFAX: (703)683-4109  
; TELEX: 899149  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7218 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; CLONE: pTZgpt-F15  
US-08-232-463-14

Query Match 6.5%; Score 74.4; DB 1; Length 7218;



;; TITLE OF INVENTION: Identification of a region of the major surface  
;; FILE REFERENCE: 4239-58054  
;; CURRENT APPLICATION NUMBER: US/09/762,724  
;; PRIOR FILING DATE: 2001-02-09  
;; PRIOR APPLICATION NUMBER: PCT/US99/18750  
;; PRIOR FILING DATE: 1999-08-17  
;; PRIOR APPLICATION NUMBER: US 60/096,805  
;; PRIOR FILING DATE: 1998-08-17  
;; NUMBER OF SEQ ID NOS: 26  
;; SOFTWARE: PatentIn Ver. 2.0  
;; SEQ ID NO 9  
;; LENGTH: 3081  
;; TYPE: DNA  
;; ORGANISM: Pneumocystis carinii sp. f. hominis  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: (1)..(3030)  
US-09-762-724-9

Query Match 3.7%; Score 42.6; DB 4; Length 3081;  
Best Local Similarity 46.2%; Pred. No. 0.031;  
Matches 141; Conservative 0; Mismatches 164; Indels 0; Gaps 0;  
QY 206 CTTTAAACCTCTCTGCTATTATTAACCTTAAGCGCTGAGCGCTCAAGACATAAAACCAITTA 265  
Db CCTTAACTCTGCAATCTGATTCCTCAATTTATTAATACATCTTTCTTCAAGCTCTAGATATCTTC 2624  
QY 266 CCAAAAAAATCACTCTCTCTCTGACACATATCTTTAGTGATATATGACATGCGCTTCG 325  
Db CAAATACCTCTGCTACCAATCAAAATCAATGCGCTTAGCTTCAGATCTGTTACGGATACATCT 2564  
QY 326 TTGCTGCTGACATTTGATTTTGTACTCAATTTGGGCTACTTTTACCTTCTGCTCAACTT 385  
Db TAACTCTCTCGTAGTATTTTATATGATTTCTGCTGATCTTGAAGTCTGTTGAACGG 2504  
QY 386 TCTCTTTGATCTCTGCGTGCATTTGCTTTTAAACGCGCTTCTCTTCTTCTTAACT 445  
Db TATCTTATCTCTGCTCTTCTGCTACTATTTTTCGATGATTAATTTTCTTCTTAACT 2444  
QY 446 CACAAAAATCACACCAATTAATCAATCAATCTGCTTCTTCAACCATATCTTCTACAC 505  
Db ATGATAAAACAAGTTGGAATTAATCAATTTCTGCTGCTTTTCTGCTGCTTTTAAAGTTTTCAA 2384  
QY 506 TTCTT 510  
Db 2383 ATGTT 2379

RESULT 5  
US-09-806-708B-22  
; Sequence 22, Application US/09806708B  
; Patent No. 6784342  
; GENERAL INFORMATION:  
; APPLICANT: The University of British Columbia  
; TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants  
; FILE REFERENCE: 4810-58741  
; CURRENT APPLICATION NUMBER: US/09/806,708B  
; CURRENT FILING DATE: 2001-04-03  
; PRIOR APPLICATION NUMBER: US 60/147,133  
; PRIOR FILING DATE: 1999-08-04  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 22  
; LENGTH: 1141  
; TYPE: DNA  
; ORGANISM: Artificial sequence  
; FEATURE:  
; NAME/KEY: promoter  
; LOCATION: (1)..(1141)  
; OTHER INFORMATION: consensus sequence of A.t., L.a., and B.n. FAEl promoters  
US-09-806-708B-22

Query Match 3.6%; Score 41.6; DB 4; Length 1141;  
Best Local Similarity 9.5%; Pred. No. 0.036;  
Matches 64; Conservative 278; Mismatches 327; Indels 8; Gaps 2;  
QY 472 AATTCACTTGTCTCTTAACCATATCTTCTACACTTCTTGTGTTATCCAAATGAAACAGAA 531  
Db 70 RWNKNSRWKMGYKKWYBCANNYSBRYHARRWKDMKTAYBMTMTNKGKTKGWRHRYWR 129  
QY 532 TCTCCCTCTTCTACTTCAAAAGTCCGACGCAAGTCAAGTATGTGTTATCGATCATCTGC 591  
Db 130 WRAMBDTVDHVYVAMNNNAWTTMCMDDKDKTRWKKNNNATGDDDDTKHMMWNGC 189  
QY 592 GCGTCGGTAGCTCAGCTGTTATTCTCTGCTGCTTTCTTTAAACAGATTACCGTTTCGAA 651  
Db 190 BTVTWVRYKTDWDWSBKRMNYGMBWKNWSYDVTYVWVWDDMCKRKRVRVTRGRMR 249  
QY 652 AAGATTTAAAGAAATACACATTTCAAGGCTATTTTAGACATGCGCCACATATCCG----- 705  
Db 250 NYMVAWBTARRRYNNGMTBAMAYRRTWNNNNNAKAMCKRAKYGMWRABVNSTCTTW 309  
QY 706 TCTATGGTAGCTACTTGTGTAGTTGTGGTAGGACTTTTGGGAAGTGGTGGTGAAGAAAG 765  
Db 310 KSKTKYRTSCHANNCRAGDANKDKHWKWSAAMGVYNNNNNNNNWTKYKARHBARWDVW 369  
QY 766 CTGAGTACAGAAATGAAGAGTTTCAACTPAGGAAAGCTCATACATTTTGTATAAACATC 825  
Db 370 WSAWKWHAANAHHYSRKKWTBYKRKTWNNNNNGTTMVKRWAWYKMDMDWBGTYNNNN 429  
QY 826 GGTTCACAGATATCATGCAAGCTTGTGTTGATTTGAAGTGTGTTGTTGATTTATCGAAGTT 885  
Db 430 NGRTYYGWTGKNQWMTYKWKANNCKWRADHKTCTHNNNTTWKMKTYNNNCYWKSMTN 489  
QY 886 TCATCGCTTTTTCCTCAATGCTATAAGCACTCTTGTGTTTACCAGTGTGCTGCTTCTTGT 945  
Db 490 GSHRBAAAVTYWVWRRYAHANNNDWYWKACTYKYBVCKWNNNYAAWYTKSSW 549  
QY 946 GTTGTCTTCTTCGCTGATGAGTGAATCAAGTTGGTTGCAATGTTTGGCCATC 1005  
Db 550 NYTSRYRWKTNWSRSDTRSMGRANNYARABHYGKWKWTRWBSHTWBHBRAGAAH 609  
QY 1006 TGGGATTTGTTCTTATGTTATCAGCATTTATGTCATGATAGAA--AGCCAGAGAAG 1063  
Db 610 YMMWYBAKCHMKAWYKAKYAGAGGSSNNNNNNNNNNNNNNNNNNATCARDYYAASRWY 669  
QY 1064 ACCAAGAGCTTCTCAGTCAAGTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1123  
Db 670 AMANAKWYKYBAANNAYYTHANNWGWGNNATDTRTWKNNNNNNNNNNNNNNNNNNNAK 729  
QY 1124 TCCAAAGCTTAGGCAAG 1140  
Db 730 NASAANKYAAAAVKA 746  
RESULT 6  
US-09-806-708B-22/c  
; Sequence 22, Application US/09806708B  
; Patent No. 6784342  
; GENERAL INFORMATION:  
; APPLICANT: The University of British Columbia  
; TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants  
; FILE REFERENCE: 4810-58741  
; CURRENT APPLICATION NUMBER: US/09/806,708B  
; CURRENT FILING DATE: 2001-04-03  
; PRIOR APPLICATION NUMBER: US 60/147,133  
; PRIOR FILING DATE: 1999-08-04  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 22  
; LENGTH: 1141  
; TYPE: DNA  
; ORGANISM: Artificial sequence  
; FEATURE:  
; NAME/KEY: promoter









APPLICANT: Drmanac, Radoje T.  
TITLE OF INVENTION: No. 6569662e1 Nucleic Acids and  
FILE REFERENCE: Polypeptides  
CURRENT APPLICATION NUMBER: US/09/620,312D  
CURRENT FILING DATE: 2000-07-19  
PRIOR APPLICATION NUMBER: 09/552,317  
PRIOR FILING DATE: 2000-04-25  
PRIOR APPLICATION NUMBER: 09/488,725  
PRIOR FILING DATE: 2000-01-21  
NUMBER OF SEQ ID NOS: 1105  
SOFTWARE: Pf\_Fl\_genes Version 1.0  
SEQ ID NO 153  
LENGTH: 5092  
TYPE: DNA  
ORGANISM: Homo sapiens  
NAME/KEY: CDS  
LOCATION: (455)..(4195)  
US-09-620-312D-153

Query Match 3.2%; Score 36.8; DB 4; Length 5092;  
Best Local Similarity 48.8%; Pred. No. 2;  
Matches 157; Conservative 0; Mismatches 162; Indels 3; Gaps 2;  
QY 782 AAGAGTTTCAACTAGGAAAGCTCATACATTTTGATAAATCGGTTCAACGATATCAT 841  
DB 4523 AAGTTTACATGCTTAAACACAGCTACCTTTGTAATACTTCATCTGATCAGAAGTGT 4582  
QY 842 GCGAAGCTTTGATGAGGAGTGGTTGTTGATTCGAGTTTCATCGCTTTTTCCTCA 901  
DB 4583 GTCATGCTTTTGA--GATGGAGTTGCTGCATTTTAGGACTATTGATACCTTTTAA 4640  
QY 902 ATGCTAAGCACTCTTTGTTTACCAGTTGCTGCTGTTCTGTTGTTCTTCTCCGTG 961  
DB 4641 TTGTTTATAATTAATTTGAAGAGAGAGACCTTCTCTCTACTCTTTCATAGAC 4700  
QY 962 ATGAG-ATGAGTGAATCAAGTTGGTTGCAATGTTTTGGCCATCTGGGATTTGTTCT 1020  
DB 4701 TGAAGTTGAATATGAATAGGCTTAACCATCATGTTGACTCTCTGTCAGAATTTAG 4760  
QY 1021 TATGTTATCAGCATTTATGTAATGATAGAACCCAGAGAACCAAGCTTCTCTAG 1080  
DB 4761 GTTGGAAATTTGTTTATTTCTTTTCATGTAATGCTTATTTGAACAGATCACTTAA 4820  
QY 1081 TCTAAAGAGAACAGAACAAA 1102  
DB 4821 GCITTAGAAGATGATTCAA 4842

RESULT 15  
US-10-204-708-45  
Sequence 45, Application US/10204708  
Patent No. 6677731  
GENERAL INFORMATION:  
APPLICANT: OLEK, Alexander  
APPLICANT: PIEPENBROCK, Christian  
APPLICANT: BERLIN, Kurt  
TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication  
FILE REFERENCE: 5013.1012  
CURRENT APPLICATION NUMBER: US/10/204,708  
CURRENT FILING DATE: 2003-05-06  
PRIOR APPLICATION NUMBER: PCT/EP01/03971  
PRIOR FILING DATE: 2001-04-06  
PRIOR APPLICATION NUMBER: DE 10019058.8  
PRIOR FILING DATE: 2000-04-06  
PRIOR APPLICATION NUMBER: DE 10019173.8  
PRIOR FILING DATE: 2000-04-07  
PRIOR APPLICATION NUMBER: DE 10032529.7  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: DE 10043826.1  
PRIOR FILING DATE: 2000-09-01

NUMBER OF SEQ ID NOS: 98  
SEQ ID NO 45  
LENGTH: 19233  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
NAME/KEY: unsure  
LOCATION: (76, 178..179, 273, 586, 648, 651, 920, 1014, 1173, 1197, 1228)  
OTHER INFORMATION: n is a or g or c or t  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (1264, 1267, 1559, 1607, 1677, 1703, 1742, 1848, 1850, 2033)  
OTHER INFORMATION: n is a or g or c or t  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (2168, 2254, 2257, 2280, 2291, 2723, 3018, 3080, 4205, 4209)  
OTHER INFORMATION: n is a or g or c or t  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (4216, 4225, 4239, 4246, 4254, 4265, 4270, 4274, 4288, 4294..4295)  
OTHER INFORMATION: n is a or g or c or t  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (4298, 4363, 4610, 4627, 4675, 4693, 4698, 4738, 4777, 4781)  
OTHER INFORMATION: n is a or g or c or t  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (4787, 4825, 4829, 4838, 4855, 5462, 5494, 5496, 5500, 5674)  
OTHER INFORMATION: n is a or g or c or t  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (5690, 5697, 5705, 5715, 5890, 5910, 5913, 6027, 6080, 6109)  
OTHER INFORMATION: n is a or g or c or t  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (6244, 6288, 6776, 6873, 6891, 6980, 7028, 7656, 7876, 8030)  
OTHER INFORMATION: n is a or g or c or t  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (8059, 8075, 8632, 8636, 9700, 9717, 9815, 9818, 9867, 9879)  
OTHER INFORMATION: n is a or g or c or t  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (10005, 10075, 10150, 10159, 10165, 10356, 11123, 11166, 11360)  
OTHER INFORMATION: n is a or g or c or t  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (11703, 12050, 12340, 12981, 13564, 13764..13765, 13787)  
OTHER INFORMATION: n is a or g or c or t  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (13846, 14168, 16566, 17006, 17217, 17650, 17653, 17656, 17898)  
OTHER INFORMATION: n is a or g or c or t  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (18130, 18143, 18151, 18188, 18498, 18549, 18936, 18946, 19165)  
OTHER INFORMATION: n is a or g or c or t  
FEATURE:  
LOCATION: (19200)  
NAME/KEY: unsure  
OTHER INFORMATION: n is a or g or c or t  
US-10-204-708-45

Query Match 3.2%; Score 36.6; DB 4; Length 19233;  
Best Local Similarity 50.3%; Pred. No. 4.7;  
Matches 90; Conservative 0; Mismatches 89; Indels 0; Gaps 0;  
QY 849 TTGTTTGAATGGAAGTGTGTTTTCATATCAAGTTTCATCGCTTTTTCATATCAT 908  
DB 3477 TTTTAAGTTGGAGTGTAGTGTGTTTTCATATGTTTGTGTTTTCATATGTTTTCG 3536

Qy	909	AAGCACTCTTTGTTTACCAAGTTGCGCTGTTCTTCTGCTGTGTCCTTCTTCCGTGATGAGAT	968
Db	3537		
Qy	969	GAGTGAATCAAAGTTGGTTGCAATGTTTTTGGCCATCTGGGGAATTTGTTCTTATGGTT	1027
Db	3597		

Search completed: November 1, 2004, 22:40:43  
Job time : 108.63 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 1, 2004, 02:40:28 ; Search time 601.255 Seconds  
(without alignments)  
9996.736 Million cell updates/sec

Title: US-09-913-767-3

Perfect score: 1145

Sequence: 1 ctgtttcatgtgttgatggt.....caagtttaggcaagatcca 1145

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

N\_Geneseq\_23Sep04:\*  
1: Geneseqn1980s:\*  
2: Geneseqn1990s:\*  
3: Geneseqn2000s:\*  
4: Geneseqn2001as:\*  
5: Geneseqn2001bs:\*  
6: Geneseqn2002as:\*  
7: Geneseqn2002bs:\*  
8: Geneseqn2003as:\*  
9: Geneseqn2003bs:\*  
10: Geneseqn2003cs:\*  
11: Geneseqn2003ds:\*  
12: Geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1145	100.0	1145	3	Aaa97921 A. thalia
2	1124.2	98.2	1292	3	Aac51119 Arabidops
3	1122.6	98.0	1295	3	Aac37289 Arabidops
4	318.4	27.8	3387	6	Abz14407 Arabidops
5	318.4	27.8	3387	8	Ada68246 Arabidops
6	317.2	27.7	1194	3	Aaa97923 A. thalia
7	308.6	27.0	1071	3	Aaa97925 A. thalia
8	304.8	26.6	2860	3	Aac47974 Arabidops
9	298.4	26.1	1047	6	Adg88190 A. thalia
10	298.4	26.1	1081	3	Aaa97924 A. thalia
11	294.4	25.7	1293	3	Aaa97922 A. thalia
12	243.2	21.2	83698	6	Abn85767 Arabidops
13	235.8	20.6	1403	3	Aac37652 Arabidops
14	190.2	16.6	1152	8	Ada70213 Rice gene
15	162.2	14.2	2175	8	Ada70748 Rice gene
16	154.6	13.5	592	10	Abx56692 Arabidops
17	130.6	11.4	1155	8	Ada69686 Rice gene
18	125.4	11.0	1040	3	Ada69568 Rice gene
19	112	9.8	1049	3	Aaa97920 A. thalia
20	104.4	9.1	332	12	Adp92185 Cotton ex
21	95.8	8.4	1225	3	Aaa97919 A. thalia

22	94.8	8.3	1173	12	ADN72520
23	92.8	8.1	2586	3	AAC44184
24	77.6	6.8	477	3	AAC37273
25	76.2	6.7	774	6	ABN99127
26	68.6	6.0	492	3	AAC36845
27	52.8	4.6	386	6	ABQ85215
28	50.4	4.4	493	3	AAC36831
29	49	4.3	13573	6	ABL33869
30	46.4	4.1	446	3	AAC37188
31	45	3.9	3090	3	Aaz94065
32	44.6	3.9	17213	6	ABL33483
33	44	3.8	15373	6	ABL32467
34	43	3.8	6145	4	AAS46547
35	42.6	3.7	3081	3	Aaz94067
36	42.4	3.7	2322	5	ADM19363
37	42	3.7	402	6	ABZ08435
38	41.8	3.7	94400	12	ADP08387
39	41.6	3.6	476	5	ADL37109
40	41.6	3.6	476	5	ADI71958
41	41.6	3.6	2000	8	ADA71938
42	41.6	3.6	2185	6	AAI45822
43	41.4	3.6	1418	3	AAC46395
44	41.2	3.6	3628	8	ABZ10117
45	40.6	3.5	555	4	ABA59777

## ALIGNMENTS

### RESULT 1

AAA97921  
ID AAA97921 standard; DNA; 1145 BP.

AC AAA97921;

DT 19-JAN-2001 (first entry)

DE A. thaliana PUP1 DNA #3.

KW PUP1; transgenic plant; nucleobase transporter; apical dominance;  
KW flowering behaviour; senescence; pesticide distribution; ds.

OS Arabidopsis thaliana.

PN DE19907209-A1.

PD 24-AUG-2000.

PF 19-FEB-1999; 99DE-01007209.

PR 19-FEB-1999; 99DE-01007209.

XX (FROM/) FROMMER W.

PI Gillissen B, Buerkle L, Andre B, Frommer WB;

XX WPI; 2000-566202/53.

PT Nucleic acid, useful for producing transgenic plants with altered  
PT nucleobase transporter, encodes a nucleobase transporter protein of  
PT Arabidopsis thaliana.

XX Claim 1f; Page 12-13; 24pp; German.

CC This invention describes a novel nucleic acid encoding a plant nucleobase  
CC transporter (I). (I) is produced by complementation of a nucleobase  
CC transport (NBT)-defective host cell with a plant gene bank by selection  
CC of NBT-positive cells. (I) is used to isolate homologous sequences from  
CC bacteria, fungi, plants, animals and humans, for expression of the  
CC encoded protein (II) in prokaryotic or eukaryotic cells, for inhibiting  
CC expression of (II) (when in antisense orientation), and to produce  
CC transgenic crop plants. The transgenic plants have modified nucleobase  
CC transport properties, e.g. altered affinity and substrate specificity



PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143342P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 21-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149360P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149428P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154035P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156598P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157533P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 18-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 22-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 25-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161922P.



PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 28-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0142878P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.

PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161992P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.

Query Match

98.0%; Score 1122.6; DB 3; Length 1295;

Best Local Similarity 98.8%; Pred. No. 4.2e-295;			
Matches 1131; Conservative 0; Mismatches 14; Indels 0; Gaps 0;			
Qy	1	CTGTTTCATGTTGATGGTATAGACCTCGAAGGGAAATTTTCAACAGAGAGAGAGTCAAC	60
Db	53	CTCCATCTTCACGTAATGGCGAACCTCGAAGGGAAATTTTCAACAGAGAGAGAGTCAAC	112
Qy	61	AAGTACTCTTGGAGGTTAAGAGTGTCTCTATGTCACTCTCCTCTTAGCTGGAGAGACA	120
Db	113	AAGTACTCTTGGAGGTTAAGAGTGTCTCTATGTCACTCTCCTCTTAGCTGGAGAGACA	172
Qy	121	ATAGCACTCTCTTAGGTAGACTTTTACTACCAAAAAGCGGTAAAGCACATGCTCGAA	180
Db	173	ATAGCACTCTCTTAGGTAGACTTTTACTACCAAAAAGCGGTAAAGCACATGCTCGAA	232
Qy	181	ACCTTGGTTAGCTTGTAGGTTTCTTTAAACCCCTTCTTGTCTATTTACTTAAAGCCT	240
Db	233	ACCTTGGTTAGCTTGTAGGTTTCTTTAAACCCCTTCTTGTCTATTTACTTAAAGCCT	292
Qy	241	GAGCGTCCAGACTAAACCAATTAACCAAAAAAATCTTCTTCTTGTGACACTATCT	300
Db	293	GAGCGTCCAGACTAAACCAATTAACCAAAAAAATCTTCTTCTTGTGACACTATCT	352
Qy	301	TTAGTGTATATTGGACTTGGCTTGTCTTCTGACATTTGATTTTGTACTCATTTGGG	360
Db	353	TTAGTGTATATTGGACTTGGCTTGTCTTCTGACATTTGATTTTGTACTCATTTGGG	412
Qy	361	CTACTTTACCTTCTGCTCAACTTCTTCTTGTATCTCTGCGTCGCAATTGGCTTTTAAAC	420
Db	413	CTACTTTACCTTCTGCTCAACTTCTTCTTGTATCTCTGCGTCGCAATTGGCTTTTAAAC	472
Qy	421	GCGTCTTCTTACTTCTTAACTTCAACAAAAATCACACCAATTTATCTCAATTCACTT	480
Db	473	GCGTCTTCTTACTTCTTAACTTCAACAAAAATCACACCAATTTATCTCAATTCACTT	532
Qy	481	GTTCTCTTAAACATATCTTCTACACTTCTTGTATCCAAATGAACAGATCTCCCTCT	540
Db	533	GTTCTCTTAAACATATCTTCTACACTTCTTGTATCCAAATGAACAGATCTCCCTCT	592
Qy	541	TCTACTTCAAAAGTCGCGACCAAGTCCAAAGTATGTGATTTGGATACATCTGCGCGT	600
Db	593	TCTACTTCAAAAGTCGCGACCAAGTCCAAAGTATGTGATTTGGATACATCTGCGCGT	652
Qy	601	AGCTCAGCTGTTTATCTCTGCTGCTTCTTAAACAGATTACGGTTCGAAAGATCTA	660
Db	653	AGCTCAGCTGTTTATCTCTGCTGCTTCTTAAACAGATTACGGTTCGAAAGATCTA	712
Qy	661	AAGAAATACATTCAGGCTATTTTACACATGGCCACATATCCGTCTATGGTAGTACT	720
Db	713	AAGAAATACATTCAGGCTATTTTACACATGGCCACATATCCGTCTATGGTAGTACT	772
Qy	721	TGTGTAGTTGTGTAGGACTTTTGGAAAGTGGTGGTGGAAAAAGCTGAGTACAGAAATG	780
Db	773	TGTGTAGTTGTGTAGGACTTTTGGAAAGTGGTGGTGGAAAAAGCTGAGTACAGAAATG	832
Qy	781	GAAAGTTTCAACTAGGAAAAAGCTACATACATTTTGTATTAACATCGGTTCAACATATCA	840
Db	833	GAAAGTTTCAACTAGGAAAAAGCTACATACATTTTGTATTAACATCGGTTCAACATATCA	892
Qy	841	TGGCAAGCTTGTGTTGATGGAAGTGTGTTGTTGATTTATCGAAGTTTCACTGCTTTTTC	900
Db	893	TGGCAAGCTTGTGTTGATGGAAGTGTGTTGTTGATTTATCGAAGTTTCACTGCTTTTTC	952
Qy	901	AATGTCATPAGCACTTTTGTGTTTACCAGTTGTGCTGTCTTGTGTTGTCTTCTTCGCT	960
Db	953	AATGTCATPAGCACTTTTGTGTTTACCAGTTGTGCTGTCTTGTGTTGTCTTCTTCGCT	1012
Qy	961	GATGAGATGAGTGGATCAAGTTGGTTGCAATGTTTTGGCCATCTGGGATTTGTTTCT	1020
Db	1013	GATGAGATGAGTGGATCAAGTTGATTTGCAATGTTTTGGCCATCTGGGATTTGTTTCT	1072
Qy	1021	TATGGTTATCAGCATTTATGTCAATGATAGAAGCCAGAAAGACCAAGAGCTTCCTCAG	1080

Db	1073	TATGTTTATCAGCATTTATGTCAATGATAGAAAGCCAGAAAGACCAAGAGTCTTCTCAG	1132
Qy	1081	TCTAAGAAGAAGAAGAACAAAAACAAGTAGATACCAATTCATGTCCAAAGCTTAGGCAAG	1140
Db	1133	TCTAAGAAGAAGAAGAACAAAAACAAGTAGATACCAATTCATGTCCAAAGCTTAGGCAAG	1192
Qy	1141	ATCCA 1145	
Db	1193	ATCCA 1197	
RESULT 4			
ID	ABZ14407	standard; DNA; 3387 BP.	
AC	ABZ14407;		
XX	21-JAN-2003	(first entry)	
DE	Arabidopsis thaliana	stress regulated gene SEQ ID NO 2212.	
XX	Arabidopsis thaliana;	plant; gene; stress; transgenic; ds.	
OS	Arabidopsis thaliana.		
PN	WO200216655-A2.		
XX	28-FEB-2002.		
XX	24-AUG-2001;	2001WO-US026685.	
PR	24-AUG-2000;	2000US-0227866P.	
PR	26-JAN-2001;	2001US-0264647P.	
PR	22-JUN-2001;	2001US-0300111P.	
XX	(SCRI )	SCRIPPS RES INST.	
PA	(SYGN )	SYNGENTA PARTICIPATIONS AG.	
XX	Harper JF,	Kreps J, Wang X, Zhu T;	
FI	WPI;	2002-304127/34.	
DR	Identifying a stress condition to which a plant cell has been exposed and		
PT	producing plants with increased tolerance to these abiotic stresses.		
XX	Claim 14;	SEQ ID NO 2212; 577pp + Sequence Listing; English.	
XX	The invention relates to identifying a stress condition to which a plant		
CC	cell has been exposed, comprising: (a) contacting nucleic acid		
CC	representative of expressed polynucleotides in the plant cell with an		
CC	array or probes representative of the plant cell genome; and (b)		
CC	detecting a profile of expressed polynucleotides in the plant cell		
CC	characteristic of a stress response. The method is useful in the		
CC	production of transgenic plants, cells and seeds and in producing plants		
CC	with increased tolerance to abiotic stress. The present sequence is that		
CC	of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used		
CC	in methods of the invention. Note: The sequence data for this patent is		
CC	not represented in the printed specification but is based on sequence		
CC	information supplied to Derwent by the European Patent Office		
XX	Sequence 3387 BP;	870 A; 781 C; 654 G; 1082 T; 0 U; 0 Other;	
Query Match 27.8%; Score 318.4; DB 6; Length 3387;			
Best Local Similarity 57.8%; Pred. No. 5.8e-76;			
Matches 593; Conservative 0; Mismatches 421; Indels 12; Gaps 1;			
Qy	28	GAAGGGAAATTTTCAACAGAGAGAGAAAGTACCAAGTACTCTTGGAGGTTAAGAGTGTCT	87
Db	2323	GAACTGAATCATTTTCAGTACCTCAACAGAGACTGTAAAGGTGGCTCCGTGTCTCC	2382
Qy	88	CTCTATGCTACTCTCTCTTTAGCTGGAGAGACAATAGCCACTCTCTTTAGGTAGACTTTAC	147
Db	2383	ATATACGAATCTTTGTCTATCTTCTGCAACCACTTCTGTACAGTTCTCTGGGTAGACTGTAC	2442







Db 723 AGCAGTCACGAGCTGGCCATTTACCAGTCTCTAGTTGCGAGTTGTGTAGTTCTCATAGG 782  
QY 738 ACTTTTGGAAAGTGGTGGGAAAGAGCTGAGTACAGAAATGGAAGAGTTTCAACTAGG 797  
Db 783 ACTTTTGGCAAGTGGAGAGTGGGAAACTTTGCCAAGTGAGATGAGAACTACAAACTCG 842  
QY 798 GAAAGCTCATACATTTTGTATAAATCGGTTCAACGATATCATGCGAAAGCTTTGTTGAT 857  
Db 843 GAAAGTGTATATGTTTGAATTTAGCTTTCAGCTCGGAGCTATTTCTCGCAAGTCTACACTCT 902  
QY 858 TGGAGTGTGGTTGATTTATCAAGTTTCATCGCTTTTTCCTCAATGTCATAGCACTCT 917  
Db 903 TGGTCTGTGGGATGATCTTCAGTCATCCTCTGTGTTCTTCCAAATTCATACAGCTGT 962  
QY 918 TTGTTTACCAGTTGTGCTCTTCTGCTGTGTGTTCTTTCGCGATGAGATGAGTGAAT 977  
Db 963 GGGATTCCTATAGTTCCAGTTCGCGAGTATGATTTTCCATGATAGATGAGCGATC 1022  
QY 978 CAAGTTGGTTGCAATGTTTGGCCATCTGGGAGTTGTTTCTTATGTTATCAGCAATTA 1037  
Db 1023 AAAAATCTTCTCCATATTTTAGCTATCTGCGGCTTCTTTCATTCGTCTATCAGCACTA 1082  
QY 1038 TGTCAATGATAGAAAG 1053  
Db 1083 CCTCGAGAAAGAAAG 1098

## RESULT 7

AAA97925  
ID AAA97925 standard; DNA; 1071 BP.  
XX  
AC AAA97925;  
XX  
DT 19-JAN-2001 (first entry)  
XX  
DE A. thaliana PUP1 DNA #7.  
XX  
KW PUP1; transgenic plant; nucleobase transporter; apical dominance;  
KW flowering behaviour; senescence; pesticide distribution; ds.  
XX  
OS Arabidopsis thaliana.  
XX  
PN DE19907209-Al.  
XX  
PD 24-AUG-2000.  
XX  
PF 19-FEB-1999; 99DE-01007209.  
XX  
PR 19-FEB-1999; 99DE-01007209.  
XX  
PA (FROM/) FROMMER W.  
XX  
PI Gillissen B, Buerkle L, Andre B, Frommer WB;  
XX  
DR WPI; 2000-566202/53.  
XX  
XX Nucleic acid, useful for producing transgenic plants with altered  
PT nucleobase transport, encodes a nucleobase transporter protein of  
PT Arabidopsis thaliana.  
XX  
PS Claim 2; Page 15; 24pp; German.  
XX  
CC This invention describes a novel nucleic acid encoding a plant nucleobase  
CC transporter (I). (I) is produced by complementation of a nucleobase  
CC transport (NBT)-defective host cell with a plant gene bank by selection  
CC of NBT-positive cells. (I) is used to isolate homologous sequences from  
CC bacteria, fungi, plants, animals and humans, for expression of the  
CC encoded protein (II) in prokaryotic or eukaryotic cells, for inhibiting  
CC expression of (II) (when in antisense orientation), and to produce  
CC transgenic crop plants. The transgenic plants have modified nucleobase  
CC transport properties, e.g. altered affinity and substrate specificity  
CC that may result in more efficient nucleobase transport in leaves, changes

CC in apical dominance, flowering behaviour and senescence, or improved  
CC distribution of pesticides. This sequence encodes the Arabidopsis  
CC thaliana PUP1 protein which is described in the method of the invention  
XX  
SQ Sequence 1071 BP; 307 A; 236 C; 208 G; 320 T; 0 U; 0 Other;  
Query Match 27.0%; Score 308.6; DB 3; Length 1071;  
Best Local Similarity 58.3%; Pred. No. 1.7e-73;  
Matches 589; Conservative 0; Mismatches 404; Indels 18; Gaps 2;  
QY 111 TGGAGAGCAATAGCAGCTCTCTTAGGTAGACTTTTACTAGAAAGGCGGTAAAGACAC 170  
Db 51 TGGCAATCAGTGTGTACAAATCTCGGCAGACTATATGAAATGGAGGAAACAGCAA 110  
QY 171 ATGGCTCGAAACCTTGGTTTCACTTGTAGGGTTTCTTTAAACCTTCTTGTCTATTATTA 230  
Db 111 ATGGCTAGCAACGGTAGTTTCACTTGTAGGCTTCTTATCTTCTTCCATATCACTCTT 170  
QY 231 CTTAAAGCCTGAGCGCTCAAGACTTAAACCAATACCAAAAAAATCTATCTTCTCTTCT 290  
Db 171 GTCAGTCAAAACACATACAACTCAGAGAGATGGCAAAATTA-----CCTCACTTAG 224  
QY 291 GACACTATCTTTAGTGTATATTTGGACTTGGCTTGGCTTGTCTGGACATTTGTTTGA 350  
Db 225 GAAACCGTCATTAGTTTACATAGTCTTGGACTTCTTGTAGGAGCAGCTTGTACCTATA 284  
QY 351 CTCATTTGGGCTACTTTTACCTTCTCTGCTCAACTTTTCTTTTGATCTCTGCGTCGCAAT 410  
Db 285 TTCCATTTGGAGCTGCTTTACCTACCTGTTTCTACCTTCTTCCCTGATCTGTCATCAGAT 344  
QY 411 GGCCTTTAAAGCGCTTCTTCTTACTTCTTAACTCAAAAAAATCAACCAATTTATATCT 470  
Db 345 AGCCTTCCACGCTTCTTCTTCTTATTTACTCAACTCACAAAAAATTTACTCTCTATCATTT 404  
QY 471 CAATTTCACTTGTCTTTAAACCAATCTTCTACACTTCTTGTATTTATCCACATGAACCA 530  
Db 405 GAAATTTCTTTTCT 462  
QY 531 ATCTCCCTCTTCT 590  
Db 463 -----TCAGATTTCCAAAAAGTTTACAAAGAGAGATGATGCAAGGTTTCGTATG 512  
QY 591 CGCGTCCGTAGCTCAGCTGGTTATTTCTTGGTGGCTTTCTTTTAAACAGATTTACCGTTCA 650  
Db 513 CACCGTTGGTGCATCTGCTGGGTTTGGTCTACTCTTATCCCTACAACAGCTAGCCTTTCG 572  
QY 651 AAAGATTTAAAGAAATACACATTTCAAGCTATTTTAGACATGGCCACATATCCGCTCTAT 710  
Db 573 TAAAGTTTTTAAAGAAAGCAAACTTTCTCAGAAAGTTATTAATATGATAATCTACATGAGTCT 632  
QY 711 GGTAGCTACTTGTGTAGTTGTGTAGGACTTTTTCGAAAGTGGTGGGAAAAAGCTGAG 770  
Db 633 AGTGGCCAGTTGTGTAGGTTGGTGGGCTTTTCTGTAGTAGGAGTGGAAAACTTTGAG 692  
QY 771 TACAGAAATGGAAGAGTTTCAACTAGGAAAGCTCATACATTTTGTATAAATCATCGTTTC 830  
Db 693 CAGTGAATGGAAAACTACAACTTGGGAAGGTATCTCTATGTCATGAACCTAGTGGGAC 752  
QY 831 AACGATATATGGCAAGCTTGTTCATTTGGAAGTGTGTTGTTGATATTCGAAGTTTCATC 890  
Db 753 AGCTGTACCTGGCAGGTAATCTCCATCGTTTGCACAGGACTGATCTTTCGAGCTTTCTCTC 812  
QY 891 GCTTTTTTCCAAATGTCATAAGCACTCTTTTGTACAGTTTGGCTGTCTTCTGCTGTTCT 950  
Db 813 CCTATTCTCAATGCAATAGCGCTTTGGGACTCCCGTGGTTCTATCTCTGGCTGTCTAT 872  
QY 951 CTTCTCCGTGATGAGATGAGTGGAAATCAAGTTGGTTCGAATGTTTTTGGCCATCTGGGG 1010  
Db 873 CATTTTCCATGACAAAAATGAACGGCTTAAAGGTGATTTCTATGATTTCTAGCTATTTGGGG 932  
QY 1011 ATTTGTTTCTTATGTTTATCAGCATTTATGTCATATGATAGAAAGCCAGGAAGACCAAGA 1070  
Db 933 TTTCTGATCTCTATGTCTTACCAACAATATCTTGTATGAAACAACTTTGAAAGAAAGTAATGA 992

QY 1071 GCTTCCTCAGTCTAAGAGAGAGAGACAAACAAAGTAGATACCAATCA 1121  
Db 993 AATACCAACAACAGATCCCTGACCGACCAGAGAGAGGGTCAAGTGA 1043

RESULT 8  
AAC47974  
ID AAC47974 standard; DNA; 2660 BP.  
XX  
AC AAC47974;  
XX  
DT 18-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 55799.  
XX  
KW Hybridisation assay; genetic mapping; gene expression control;  
KW protein identification; signal transduction pathway; metabolic pathway;  
KW promoter; termination sequence; ss.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-00301439.  
XX  
PR 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 28-APR-1999; 99US-0130891P.  
PR 30-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 04-MAY-1999; 99US-0132407P.  
PR 05-MAY-1999; 99US-0132484P.  
PR 06-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 07-MAY-1999; 99US-0132487P.  
PR 11-MAY-1999; 99US-0132863P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.

17-JUN-1999; 99US-0139492P.  
18-JUN-1999; 99US-0139454P.  
18-JUN-1999; 99US-0139455P.  
18-JUN-1999; 99US-0139456P.  
18-JUN-1999; 99US-0139457P.  
18-JUN-1999; 99US-0139458P.  
18-JUN-1999; 99US-0139459P.  
18-JUN-1999; 99US-0139460P.  
18-JUN-1999; 99US-0139461P.  
18-JUN-1999; 99US-0139462P.  
18-JUN-1999; 99US-0139463P.  
18-JUN-1999; 99US-0139750P.  
18-JUN-1999; 99US-0139763P.  
21-JUN-1999; 99US-0139817P.  
22-JUN-1999; 99US-0139899P.  
23-JUN-1999; 99US-0140353P.  
23-JUN-1999; 99US-0140354P.  
24-JUN-1999; 99US-0140695P.  
28-JUN-1999; 99US-0140823P.  
29-JUN-1999; 99US-0140991P.  
30-JUN-1999; 99US-0141287P.  
01-JUL-1999; 99US-0141842P.  
01-JUL-1999; 99US-0142154P.  
02-JUL-1999; 99US-0142055P.  
06-JUL-1999; 99US-0142390P.  
08-JUL-1999; 99US-0142803P.  
09-JUL-1999; 99US-0142920P.  
12-JUL-1999; 99US-0142977P.  
13-JUL-1999; 99US-0143542P.  
14-JUL-1999; 99US-0143624P.  
15-JUL-1999; 99US-0144005P.  
16-JUL-1999; 99US-0144085P.  
16-JUL-1999; 99US-0144086P.  
19-JUL-1999; 99US-0144325P.  
19-JUL-1999; 99US-0144331P.  
19-JUL-1999; 99US-0144332P.  
19-JUL-1999; 99US-0144333P.  
19-JUL-1999; 99US-0144334P.  
20-JUL-1999; 99US-0144352P.  
20-JUL-1999; 99US-0144632P.  
20-JUL-1999; 99US-0144884P.  
21-JUL-1999; 99US-0144814P.  
21-JUL-1999; 99US-0145086P.  
21-JUL-1999; 99US-0145088P.  
22-JUL-1999; 99US-0145085P.  
22-JUL-1999; 99US-0145087P.  
22-JUL-1999; 99US-0145089P.  
22-JUL-1999; 99US-0145192P.  
23-JUL-1999; 99US-0145145P.  
23-JUL-1999; 99US-0145218P.  
23-JUL-1999; 99US-0145224P.  
26-JUL-1999; 99US-0145276P.  
27-JUL-1999; 99US-0145913P.  
27-JUL-1999; 99US-0145918P.  
27-JUL-1999; 99US-0145919P.  
28-JUL-1999; 99US-0145951P.  
02-AUG-1999; 99US-0146386P.  
02-AUG-1999; 99US-0146388P.  
02-AUG-1999; 99US-0146389P.  
03-AUG-1999; 99US-0147038P.  
04-AUG-1999; 99US-0147204P.  
04-AUG-1999; 99US-0147302P.  
05-AUG-1999; 99US-0147192P.  
05-AUG-1999; 99US-0147260P.  
06-AUG-1999; 99US-0147303P.  
06-AUG-1999; 99US-0147416P.  
09-AUG-1999; 99US-0147493P.  
09-AUG-1999; 99US-0147935P.  
10-AUG-1999; 99US-0148171P.  
11-AUG-1999; 99US-0148319P.  
12-AUG-1999; 99US-0148341P.  
13-AUG-1999; 99US-0148565P.

PR	13-AUG-1999;	99US-0148684P.
PR	16-AUG-1999;	99US-0149368P.
PR	17-AUG-1999;	99US-0149175P.
PR	18-AUG-1999;	99US-0149426P.
PR	20-AUG-1999;	99US-0149722P.
PR	20-AUG-1999;	99US-0149723P.
PR	20-AUG-1999;	99US-0149929P.
PR	23-AUG-1999;	99US-0149902P.
PR	23-AUG-1999;	99US-0149930P.
PR	25-AUG-1999;	99US-0150566P.
PR	26-AUG-1999;	99US-0150884P.
PR	26-AUG-1999;	99US-0151065P.
PR	27-AUG-1999;	99US-0151066P.
PR	27-AUG-1999;	99US-0151080P.
PR	30-AUG-1999;	99US-0151303P.
PR	31-AUG-1999;	99US-0151438P.
PR	01-SEP-1999;	99US-0151930P.
PR	10-SEP-1999;	99US-0152363P.
PR	13-SEP-1999;	99US-0153070P.
PR	15-SEP-1999;	99US-0153758P.
PR	16-SEP-1999;	99US-0154018P.
PR	20-SEP-1999;	99US-0154039P.
PR	22-SEP-1999;	99US-0154779P.
PR	23-SEP-1999;	99US-0155139P.
PR	24-SEP-1999;	99US-0155486P.
PR	28-SEP-1999;	99US-0156458P.
PR	29-SEP-1999;	99US-0156596P.
PR	04-OCT-1999;	99US-0157117P.
PR	05-OCT-1999;	99US-0157753P.
PR	06-OCT-1999;	99US-0157865P.
PR	07-OCT-1999;	99US-0158029P.
PR	08-OCT-1999;	99US-0158232P.
PR	12-OCT-1999;	99US-0158369P.
PR	13-OCT-1999;	99US-0159293P.
PR	13-OCT-1999;	99US-0159294P.
PR	13-OCT-1999;	99US-0159295P.
PR	14-OCT-1999;	99US-0159329P.
PR	14-OCT-1999;	99US-0159330P.
PR	14-OCT-1999;	99US-0159331P.
PR	14-OCT-1999;	99US-0159637P.
PR	14-OCT-1999;	99US-0159638P.
PR	18-OCT-1999;	99US-0159584P.
PR	21-OCT-1999;	99US-0160741P.
PR	21-OCT-1999;	99US-0160767P.
PR	21-OCT-1999;	99US-0160768P.
PR	21-OCT-1999;	99US-0160770P.
PR	21-OCT-1999;	99US-0160814P.
PR	21-OCT-1999;	99US-0160815P.
PR	22-OCT-1999;	99US-0160980P.
PR	22-OCT-1999;	99US-0160981P.
PR	25-OCT-1999;	99US-0160989P.
PR	25-OCT-1999;	99US-0161404P.
PR	25-OCT-1999;	99US-0161405P.
PR	25-OCT-1999;	99US-0161406P.
PR	26-OCT-1999;	99US-0161359P.
PR	26-OCT-1999;	99US-0161360P.
PR	26-OCT-1999;	99US-0161361P.
PR	28-OCT-1999;	99US-0161920P.
PR	28-OCT-1999;	99US-0161992P.
PR	28-OCT-1999;	99US-0161993P.
PR	29-OCT-1999;	99US-0162142P.
Query Match 26.6%; Score 304.8; DB 3; Length 2660;		
Best Local Similarity 58.1%; Pred. No. 2.6e-72;		
Matches 567; Conservative 0; Mismatches 402; Indels 21; Gaps 2;		
QY	70	TGGAGGTTAAGAGTCTCTCTATGTCACCTCTCTCTTCTAGCTGGAGAGACAATAGCCACT 129
DB	191	TGTTGGGTTCTGTATCCGTTAAACATCTTTTCCCTTATTTGGTGGTCAAGCTCTTCGGTT 250
QY	130	CTCTAGGTAGACTTTACTACGAAAGGCGGTAAAGCACATGCTCGAAACCTTGTT 189

Db	251	CTTCTTGTTAGTGTCTTTACTATGATGAAGGTGGAAACAGTAATGATGGCAACTCTTGTT 310
QY	190	CAGCTTGTTAGGTTTCCTTTAAACCCTTCTCTGCTATTATTACTTAAAGCCTGAGCGCTCC 249
Db	311	CAACGGCTGCTTTTCCCTATACT-----CTATATCCCGCTTTTGTCTACTCCGCT 361
QY	250	AAGACTAAACCATTAACCAAAAAAATACTTCTTCTCTTTGACACTATCTTTTAGTGTAT 309
Db	362	TCGGCTAGTAGAGTCTTCGGAGAGTTCTGTTTCACTCAAGTACATATGTTTGTATCTAT 421
QY	310	ATTGACTTGGCTGCTTGTGCTGGACATTTGTTTGTACTATTTTGGGGTACTTTTAC 369
Db	422	GTTTTCGCTGGTGTGATCAITGCTGGAGATAATATGTTATATCTCTGTTGGACTTTTGTAC 481
QY	370	CTTCCTGCTCAACTTCTCTTTTGATCTCTGCTCGCAATTTGGCTTTTAAAGCCCTCTTC 429
Db	482	CTCTCTGCATCGACGTATTTCGCTCATTTTGGCTACTCAGTTAGCTTTTCAACGGGTGTT 541
QY	430	TCCTACTTCTTAAACTCAAAAAAATCACACATTTATATCTCAATTTCACTTGTCTCTTA 489
Db	542	TCTTATTTTCATCAATGCTCAGAAGTTCACTGCTTTGATTTCTCAACTCCGTTGTTCTCTTG 601
QY	490	ACCATATCTTCTACACTTCTTTGTTATCCAAATGACCAAGACAGATCTCCCTCTTCTACTTCA 549
Db	602	TCGTTCTCCGCTGCTTTGATAGCTCTCAATGATGATGCGGATCTCTCTCTGTT----- 655
QY	550	AAGTCGCGAGCAAGTCCCAAGTATGTGATTCGATACATCTCGCGGTGCGTAGCTCAGCT 609
Db	656	-----GTCTCAAGTCTAAGTATATTTGTTGGTTTGTGTACACTTGTCTGCTGCTGCT 709
QY	610	GGTTATCTCTGCTGCTTTCTTTAAAGATACCGGTTTCGAAAAGATTTCTAAAGAAATAC 669
Db	710	CTCTATTTCTGTTGCTATCTCTTATGCACTTCGTTTCGAGAGATTTCTGAAGGGGAG 769
QY	670	ACATTCAGGCTATTTTAGACATGGCCACATATCGGCTATCTGTTAGTACTTGTGTAGTT 729
Db	770	ACGTTTCTGCTGCTTTGAAATGCAATCTACACTTCTTTAGTGGCGACTTGTGTTTCG 829
QY	730	GTGGTAGGACTTTTGGAAAGTGGTGGTGGAAAGCTGAGTACAGAAATGGAAGAGTTT 789
Db	830	GTTATAGGGCTTTTGTAGCGGGGAATGGAGACGCTGCATGGGAAATGGAAGGTTAT 889
QY	790	CAACTAGGGAAGCTCATACATTTTGATAAACATCGGTTTCAACGATATCATGGCAAGCT 849
Db	890	CATAAAGGGCAAGCTCTTATGTACTGACCTTGGTCTGACAGCAGTACTTTGGCAAGTG 949
QY	850	TGTTTGATTGGAAGTGTGTTTGAATATCGAAGTTTTCATCGCTTTTTCCTCAATGTCTATA 909
Db	950	TGTTCTGTTGGAGTCGTGGGTTTGATATTTCTGGTGCAGCTCGCTCTTCTCAACGCTCAT 1009
QY	910	AGCACTCTTCTTTTACCAAGTTGTCCTGTTCTTCTGCTGTTGCTTCTTCCGTTAGAGATG 969
Db	1010	AGTAGGCTCTCTCTAGCTGTGACTCCACTAGCAGCTTTGGTTGTGTTCCGTTGATAAATG 1069
QY	970	AGTGAATCAAGTTGGTTGCAATGTTTTTGGCCCATCTGGGGAATTTGTTTCTTATGTTTAT 1029
Db	1070	AGTGGTGAAGATTATGGCAATGCTGATCCCTATTTGGGGTTTCGCTTCTTATGTTTAC 1129
QY	1030	CAGCATTTATGTCATGATAGAAAGCCAGAGAACCAAGAGCTTCTCTCA 1079
Db	1130	CAGAAATCATATTGATGACTTGAAGAGTAAGACAGCAGCACCAACCAAGCTCA 1179
RESULT 9		
ADG88190		
ID ADG88190 standard; cDNA; 1047 BP.		
XX	AC	ADG88190;
XX	AC	
DT	22-APR-2004	(first entry)
XX		
DE		A. thaliana RPP4-upregulated pathogen infection-related gene #632.
XX		



XX DE19907209-A1.  
XX 24-AUG-2000.  
XX 19-FEB-1999; 99DE-01007209.  
XX 19-FEB-1999; 99DE-01007209.  
XX (FROM/) FROMMER W.  
XX Gillissen B, Buerkle L, Andre B, Frommer WB;  
XX WPI; 2000-566202/53.  
XX Nucleic acid, useful for producing transgenic plants with altered  
PT nucleobase transport, encodes a nucleobase transporter protein of  
PT Arabidopsis thaliana.  
XX Claim 2; Page 14-15; 24pp; German.  
XX This invention describes a novel nucleic acid encoding a plant nucleobase  
CC transporter (I). (I) is produced by complementation of a nucleobase  
CC transport (NBT)-defective host cell with a plant gene bank by selection  
CC of NBT-positive cells. (I) is used to isolate homologous sequences from  
CC bacteria, fungi, plants, animals and humans, for expression of the  
CC encoded protein (II) in prokaryotic or eukaryotic cells, for inhibiting  
CC expression of (II) (when in antisense orientation), and to produce  
CC transgenic crop plants. The transgenic plants have modified nucleobase  
CC transport properties, e.g. altered affinity and substrate specificity  
CC that may result in more efficient nucleobase transport in leaves, changes  
CC in apical dominance, flowering behaviour and senescence, or improved  
CC distribution of pesticides. This sequence encodes the Arabidopsis  
CC thaliana PUP1 protein which is described in the method of the invention  
XX  
XX Sequence 1081 BP; 322 A; 231 C; 205 G; 323 T; 0 U; 0 Other;

Query Match 26.1%; Score 298.4; DB 3; Length 1081;  
Best Local Similarity 58.3%; Pred. No. 1e-70;  
Matches 571; Conservative 0; Mismatches 391; Indels 18; Gaps 2;  
QY 140 GACTTTACTACGAAAGCGGTAAAGACACATGGCTGAAACCTTGGTTCAGCTGTAG 199  
Db 89 GAGTATACATGACACGAGGAAACAGTAAATGGCTGACAAACGGTAGTTCACCTTTG 148  
QY 200 GGTTCCTTTAAACCTTCTCTGTATTATTACTTAAAGCCTGAGCGCTCAAGACTAAA 259  
Db 149 GCTTTCCTGTGCTACTTCCATATTATATCTTGTCATTTAAACACATGCAACACTGATA 208  
QY 260 CCATTACCAAAAAAATCTCTTCTCTTGACACTATCTTTAGTGATATTGGACTTG 319  
Db 209 GAGATGGAAGAAAGAACCTCACCTAG-----GAAACGGTATTGGTTACGTAGTCTTG 262  
QY 320 GCTTGGCTTGTGAGACATATGTATTTTGTACTCAATTTGGGCTACTTTACCTTCTGTCT 379  
Db 263 GACTTCTCTAGGAGCAGATTTGTATCTTCTTCCATTTGGACTTCTTTACTTACCCGTTT 322  
QY 380 CAACCTTCTCTTGTATCTGCTGGTGCATTTGGCTTTTAAACCGCTTCTCTACTTCC 439  
Db 323 CTACCTATTCCCTGATCTGTGATCTCAGTTAGCCTTCAATGCTTTCTCTCTATTTC 382  
QY 440 TAAACTCAGAAAAATTCACCACTTTATATCTCAATTCACCTTGTCTCTTAAACCATATCTT 499  
Db 383 TTAACCTCAGAAAAATTCACCCCTATCATTTTAAATCTCTTTTCTCTCACTATATCTT 442  
QY 500 CTACACTTCTGTATTCCAAACATGAAACAGAAATCTCCCTCTCTTCAAAAGTCGAG 559  
Db 443 CCACCTTACTTGCATTTCAATAATGAGGAGACAGACTCCACAAAAAGTTACAAAAGGAG--- 499  
QY 560 CCAAGTCCCAAGTATGATGATGATACATCTGCGGCTCGTAGCTCGCTGTTATCTC 619  
Db 500 -----AGTATGTCAAGGTTTCATATGACCGGTGCTGCTGCTGTTATGCTC 550

QY 620 TGGTCTCTTTTAAACAGATTACGGTTTGAAGAAATCTTAAAGAAATACACATTCAAGG 679  
Db 551 TAGTCTTATCCCTACAAACAGCTAGCTTTCTTAAAGAGTCTTAAAGAGCAAAATTTCTCAG 610  
QY 680 CTATTTTACACATGSCCACAATATCGTCTATAGGTAGCTACTTGTGTAGTTGGTAGGAC 739  
Db 611 AAGTTATGGATATGATAATCTACGTGAGTCTAGTGGCCAGTTGTCTTACCGTGGTGGGC 670  
QY 740 TTTTGGAAAGTGGTGGTGGAAAGCTGAGTACAGAAATGGAAGAGTTTCAACTAGGA 799  
Db 671 TTTTGGCTAGCAGTGGTGGAAACTTTGAGCAGTGAATGGATAACTACAAACATGGA 730  
QY 800 AAAGCTCATACATTTTGTATAAACATCGGTTCAACGATATCATGGCAAGCTTGTGTTGATG 859  
Db 731 AGGTATCTCATATTATGAACCTAGTGTGACAGCTGTTTACCTGGCAGGTATTCTCCATCG 790  
QY 860 GNAAGTGTGGTTTGNATTATCGAAGTTTTCATCGCTTTTTCATGTCATAGCACTCTTT 919  
Db 791 GTGGCACAGGACTGATCTTCGAGCTCTCTCTATCTCAATGCAATAAGCGTTTGG 850  
QY 920 GTTTACCAAGTTGTGCTGCTTCTTCTGCTGTGCTTCTTCCGTGATGAGATGAGTGAATCA 979  
Db 851 GACTCCAGTGGTTCCTATCTTGGCTGTAAATCAITTTTCATGACAAAATGAATGGTTAA 910  
QY 980 AGTTGGTTGCAATGTTTTTGGCCATCTCGGGAATTTGTTTCTTATGGTTATCAGCAATTATG 1039  
Db 911 AGGTGATTTCTATGATCTCTAGCTATTGGGGTTTTCACCTCTCTATCTTACCAACAATATC 970  
QY 1040 TCAATGATAGAAAGCCAGAAAGACCAAGAGCTTCCCTCAGTCTTAAAGAGAGAGAAC 1099  
Db 971 TTGATGACAAAAAATCTTGAAGAAAAATCATGAATCACAACACAGAAATCCCTGACCCAC 1030  
QY 1100 AAAAAACAAGTAGATACCATT 1119  
Db 1031 CAGAAGCAGAGAGTCAACT 1050  
RESULT 11  
AAA97922  
ID AAA97922 standard; DNA; 1293 BP.  
XX  
AC AAA97922;  
XX  
DT 19-JAN-2001 (first entry)  
XX  
DE A. thaliana PUP1 DNA #4.  
XX  
KW PUP1; transgenic plant; nucleobase transporter; apical dominance;  
KW flowering behaviour; senescence; pesticide distribution; ds.  
XX  
OS Arabidopsis thaliana.  
XX  
PN DE19907209-A1.  
XX  
PD 24-AUG-2000.  
XX  
PF 19-FEB-1999; 99DE-01007209.  
XX  
PR 19-FEB-1999; 99DE-01007209.  
XX  
PA (FROM/) FROMMER W.  
XX  
PI Gillissen B, Buerkle L, Andre B, Frommer WB;  
XX  
WPI; 2000-566202/53.  
XX  
PT Nucleic acid, useful for producing transgenic plants with altered  
PT nucleobase transport, encodes a nucleobase transporter protein of  
PT Arabidopsis thaliana.  
XX  
PS Claim 1f; Page 13; 24pp; German.  
XX  
CC This invention describes a novel nucleic acid encoding a plant nucleobase

CC transporter (I). (I) is produced by complementation of a nucleobase  
CC transport (NBT)-defective host cell with a plant gene bank by selection  
CC of NBT-positive cells. (I) is used to isolate homologous sequences from  
CC bacteria, fungi, plants, animals and humans, for expression of the  
CC encoded protein (ii) in prokaryotic or eukaryotic cells, for inhibiting  
CC expression of (ii) (when in antisense orientation), and to produce  
CC transgenic crop plants. The transgenic plants have modified nucleobase  
CC transport properties, e.g. altered affinity and substrate specificity  
CC that may result in more efficient nucleobase transport in leaves, changes  
CC in apical dominance, flowering behaviour and senescence, or improved  
CC distribution of pesticides. This sequence encodes the Arabidopsis  
CC thaliana PUP1 protein which is described in the method of the invention  
XX  
SQ Sequence 1293 BP; 345 A; 288 C; 247 G; 413 T; 0 U; 0 Other;

Query Match 25.7%; Score 294.4; DB 3; Length 1293;  
Best Local Similarity 56.3%; Pred. No. 1.4e-69;  
Matches 578; Conservative 0; Mismatches 436; Indels 12; Gaps 1;

QY 28 GAAGGGAATTTTCAACAGAGAGAGAGTACAGTACTCTTGAGGTTAAGAGTGTCT 87  
DB 181 GAAATCGAATCTTCTCGTACTCTCAATCGAAGAACTATAAGAAATGGCTTCGTATTTCC 240  
QY 88 CTCTATGTCACCTCTCTTCTTACGTCGAGAGACAATAGCCACTCTCTTAGGTAGACTTTAC 147  
DB 241 ATTACGTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 300  
QY 148 TAGCAAAAGGGGGTAAAGACATAGGCTCGAAACCTTGGTTCAGCTTGTAGGGTTTCC 207  
DB 301 TATGAAATATGGTGAAGAGTACATGATGGACACTTGTCCAACTAATCGGCTTCCCT 360  
QY 208 TTAACCCCTTCTGTATATATATTAAGCTGAGCGGTCCAGACTAAACCAATACC 267  
DB 361 GTTCTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 420  
QY 268 AAAAAAATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 327  
DB 421 TTCAGAAATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 480  
QY 328 GTTGTGACATATGATTTTGTATCTATTTGGGCTACTTTACCTTCTGTCCTCAACTTTC 387  
DB 481 GTGTCTGCTAATCTTATATGTCCTCTGTGGTTTACTATATCTTACCAGTTTCTACTTTC 540  
QY 388 TCTTTGATCTGCGTCGCAATGGCTTTTAACGCGCTTCTCTTACTTCTTAACTCA 447  
DB 541 TCCTCATCTTGGCTCAAAATGGCTTCTACTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 600  
QY 448 CAAAAAATCACCAATTTATCTCAATTCACCTTGTCTTCTTAAACCATATCTTCTACACTT 507  
DB 601 CAGAAATTCACACCTTCTCATTTGTGAATCTCTGTTTCTTCTTCTTCTTCTTCTTCTTCTTCT 660  
QY 508 CTTGTTATCCAAATGAAACCAATCTCCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 567  
DB 661 CTCGTGGTCAACACTGTTCGGAA-----AACACAGCAAAAGTGTCTAGAGTA 708  
QY 568 AAGTATGTATGGATATCATCTGCGCGTCGTAGCTCAGCTGGTTATCTTCTGTCGTT 627  
DB 709 AAATATGTATGGAT 768  
QY 628 TCTTTAACAAGATTACGCGTTTGAAAGATTCTTAAAGAAATACACATTTCAAGGCTATTTTA 687  
DB 769 TCCCTGGTACACTGATCTCCTCAGGAAGTTTAAAGAGCAACATCTTCAACGGTCACT 828  
QY 688 GACATGGCCACATATCCGTCTATGTAGTACTTGTGTAGTGTGGTATAGGACTTTTGGGA 747  
DB 829 GACTTGGTGGCTTACCAATCTCTAGTTCGAAGCTGTGTGGTTCTCATAGGACTTTTCGCA 888  
QY 748 AGTGTGGTGGTGAAGAGCTGATACAGAAATGGAAGTTCACACTAGGGAAGCTCA 807  
DB 889 AGCGGGAGTGGAAATCTTTTAAAGTGAATGGAAATCTCAAACTGGGGAAGTGGCA 948  
QY 808 TACATTTTGAATAACATCGGTTTCAACGATATCATGGCAAGCTTGTGTTGATTTGGAAGTGT 867

DB 949 TACGTTATGACTTTGGGCTCGATAGCTATTTCTTGGCAAGTCTACACCATTTGGCGTCTG 1008  
QY 868 GGTGTTGATTATCGAAGTTTTCATCGCTTTTTCATGTCATATAGCACTCTTTGTTTACCA 927  
DB 1009 GGACTGATCTTTGAGTCACTCTTCTGTTCTTCTCCAAATTCCTAATCTGCTGGGATTCCT 1068  
QY 928 GTTGTGCTGTTCTTCTGTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 987  
DB 1069 ATAGTTCAGTTGTAGCAGTATTTGTTTTCATGATATAAATGAACGGTCAAGATCTTC 1128  
QY 988 GCAATGTTTGGCCATCTGGGATTTGTTTCTTATGTTTATCAGCAATATGTCATATGAT 1047  
DB 1129 TCCATCATTTTAGCTATCTGGGATTCATTTCTTCTATCAGCACTACCTCGACGAA 1188  
QY 1048 AGAAAG 1053  
DB 1189 AAGAAG 1194

#### RESULT 12

ABN85767  
ID ABN85767 standard; cDNA; 83698 BP.

AC ABN85767;

DT 21-OCT-2002 (first entry)

DE Arabidopsis yellow stripe1-like 4 encoding cDNA SEQ ID NO 9.

KW Maize; transgenic; plant; yellow stripe1-like; ysl; Arabidopsis;  
KW iron uptake; bioremediation; yellow stripe 1; ysl; gene; ss.

OS Arabidopsis sp.

PN WO200240688-A2.

PD 23-MAY-2002.

PF 16-NOV-2001; 2001WO-US043101.

PR 16-NOV-2000; 2000US-0249222P.

PA (UYVA ) UNIV YALE.

PI Walker EL, Dellaporta S;

DR WPI; 2002-490144/52.

DR P-PSDB; ABB83920.

XX New yellow stripe1 and yellow stripe1-like genes, useful for altering the  
XX distribution of iron within the plant body so that edible parts of crop  
XX plants have more iron, or for producing plants useful in enhancing iron  
XX uptake from soil.

PS Claim 1; Page 114-159; 187pp; English.

XX The invention relates to an isolated nucleic acid molecule (I), maize  
XX yellow stripe 1 (ysl) or yellow stripe1-like (ysl) from Arabidopsis  
XX (ABN85763-ABN85771). (I) is useful for generating transgenic plants which  
XX can be used for enhancing iron uptake from soil and for bioremediation of  
XX metal or heavy metal contaminated soil. (I) may also be used to alter the  
XX distribution of iron within the plant body so that edible parts of crop  
XX plants have more iron. Transgenic plants may also be used in conventional  
XX plant breeding schemes to produce progeny which also contain the gene of  
XX interest. The present sequence is that of the Arabidopsis ysl encoding  
XX cDNA of the invention

SQ Sequence 83698 BP; 26295 A; 15150 C; 15104 G; 27149 T; 0 U; 0 Other;

Query Match 21.2%; Score 243.2; DB 6; Length 83698;

Best Local Similarity 54.4%; Pred. No. 5.9e-55;

Matches 547; Conservative 0; Mismatches 438; Indels 21; Gaps 2;





PR	18-JUN-1999;	99US-0139463P.	PR	26-AUG-1999;	99US-0150884P.
PR	18-JUN-1999;	99US-0139750P.	PR	27-AUG-1999;	99US-0151065P.
PR	18-JUN-1999;	99US-0139763P.	PR	27-AUG-1999;	99US-0151066P.
PR	21-JUN-1999;	99US-0139817P.	PR	27-AUG-1999;	99US-0151080P.
PR	22-JUN-1999;	99US-0139899P.	PR	30-AUG-1999;	99US-01511303P.
PR	23-JUN-1999;	99US-0140353P.	PR	31-AUG-1999;	99US-0151438P.
PR	23-JUN-1999;	99US-0140354P.	PR	01-SEP-1999;	99US-0151930P.
PR	24-JUN-1999;	99US-0140695P.	PR	07-SEP-1999;	99US-0152363P.
PR	28-JUN-1999;	99US-0140823P.	PR	10-SEP-1999;	99US-0153070P.
PR	29-JUN-1999;	99US-0140991P.	PR	13-SEP-1999;	99US-0153758P.
PR	30-JUN-1999;	99US-0141287P.	PR	15-SEP-1999;	99US-0154018P.
PR	01-JUL-1999;	99US-0141842P.	PR	16-SEP-1999;	99US-0154039P.
PR	01-JUL-1999;	99US-0142154P.	PR	20-SEP-1999;	99US-0154779P.
PR	02-JUL-1999;	99US-0142055P.	PR	22-SEP-1999;	99US-0155139P.
PR	06-JUL-1999;	99US-0142390P.	PR	23-SEP-1999;	99US-0155486P.
PR	08-JUL-1999;	99US-0142803P.	PR	24-SEP-1999;	99US-0155659P.
PR	09-JUL-1999;	99US-0142920P.	PR	28-SEP-1999;	99US-0156458P.
PR	12-JUL-1999;	99US-0142977P.	PR	29-SEP-1999;	99US-0156596P.
PR	13-JUL-1999;	99US-0143542P.	PR	04-OCT-1999;	99US-0157117P.
PR	14-JUL-1999;	99US-0143621P.	PR	05-OCT-1999;	99US-0157753P.
PR	15-JUL-1999;	99US-0144005P.	PR	06-OCT-1999;	99US-0157865P.
PR	16-JUL-1999;	99US-0144085P.	PR	07-OCT-1999;	99US-0158029P.
PR	16-JUL-1999;	99US-0144086P.	PR	08-OCT-1999;	99US-0158232P.
PR	19-JUL-1999;	99US-0144325P.	PR	12-OCT-1999;	99US-0158369P.
PR	19-JUL-1999;	99US-0144331P.	PR	13-OCT-1999;	99US-0159293P.
PR	19-JUL-1999;	99US-0144332P.	PR	13-OCT-1999;	99US-0159294P.
PR	19-JUL-1999;	99US-0144333P.	PR	13-OCT-1999;	99US-0159295P.
PR	19-JUL-1999;	99US-0144334P.	PR	14-OCT-1999;	99US-0159329P.
PR	19-JUL-1999;	99US-0144335P.	PR	14-OCT-1999;	99US-0159330P.
PR	20-JUL-1999;	99US-0144352P.	PR	14-OCT-1999;	99US-0159331P.
PR	20-JUL-1999;	99US-0144632P.	PR	14-OCT-1999;	99US-0159637P.
PR	20-JUL-1999;	99US-0144884P.	PR	14-OCT-1999;	99US-0159638P.
PR	21-JUL-1999;	99US-0144814P.	PR	18-OCT-1999;	99US-0159584P.
PR	21-JUL-1999;	99US-0145086P.	PR	21-OCT-1999;	99US-0160741P.
PR	21-JUL-1999;	99US-0145088P.	PR	21-OCT-1999;	99US-0160767P.
PR	22-JUL-1999;	99US-0145085P.	PR	21-OCT-1999;	99US-0160768P.
PR	22-JUL-1999;	99US-0145087P.	PR	21-OCT-1999;	99US-0160770P.
PR	22-JUL-1999;	99US-0145089P.	PR	21-OCT-1999;	99US-0160814P.
PR	22-JUL-1999;	99US-0145192P.	PR	22-OCT-1999;	99US-0160815P.
PR	23-JUL-1999;	99US-0145145P.	PR	22-OCT-1999;	99US-0160980P.
PR	23-JUL-1999;	99US-0145218P.	PR	22-OCT-1999;	99US-0160981P.
PR	23-JUL-1999;	99US-0145224P.	PR	22-OCT-1999;	99US-0160989P.
PR	26-JUL-1999;	99US-0145276P.	PR	25-OCT-1999;	99US-0161404P.
PR	27-JUL-1999;	99US-0145913P.	PR	25-OCT-1999;	99US-0161405P.
PR	27-JUL-1999;	99US-0145918P.	PR	25-OCT-1999;	99US-0161406P.
PR	27-JUL-1999;	99US-0147038P.	PR	25-OCT-1999;	99US-0161359P.
PR	03-AUG-1999;	99US-0147204P.	PR	26-OCT-1999;	99US-0161360P.
PR	04-AUG-1999;	99US-0147302P.	PR	26-OCT-1999;	99US-0161361P.
PR	05-AUG-1999;	99US-0147260P.	PR	26-OCT-1999;	99US-0161920P.
PR	06-AUG-1999;	99US-0147303P.	PR	28-OCT-1999;	99US-0161922P.
PR	06-AUG-1999;	99US-0147416P.	PR	28-OCT-1999;	99US-0161993P.
PR	09-AUG-1999;	99US-0147933P.	PR	29-OCT-1999;	99US-0162142P.
PR	10-AUG-1999;	99US-0148171P.			
PR	11-AUG-1999;	99US-0148319P.			
PR	12-AUG-1999;	99US-0148341P.			
PR	13-AUG-1999;	99US-0148565P.			
PR	13-AUG-1999;	99US-0148684P.			
PR	16-AUG-1999;	99US-0149368P.			
PR	17-AUG-1999;	99US-0149175P.			
PR	18-AUG-1999;	99US-0149426P.			
PR	20-AUG-1999;	99US-0149722P.			
PR	20-AUG-1999;	99US-0149723P.			
PR	20-AUG-1999;	99US-0149929P.			
PR	23-AUG-1999;	99US-0149902P.			
PR	23-AUG-1999;	99US-0149930P.			
PR	25-AUG-1999;	99US-0150566P.			
			Query Match 20.6%; Score 235.8; DB 3; Length 1403;		
			Best Local Similarity 54.1%; Pred. No. 1.3e-53;		
			Matches 545; Conservative 0; Mismatches 432; Indels 30; Gaps 2;		
Qy	69	TTGGAGGTTAAGAGTGCTCTCTATATGTCACCTCTCTTACGTGAGGAGAGACAATAGCCAC	128	TTGGAGGTTAAGAGTGCTCTCTATATGTCACCTCTCTTACGTGAGGAGAGACAATAGCCAC	128
Db	271	TTGGTGGATTCTTGCTCTTTATAAGTATCTTCTCTCATCTCTGCTCAAGCCATTGCTGT	330	TTGGTGGATTCTTGCTCTTTATAAGTATCTTCTCTCATCTCTGCTCAAGCCATTGCTGT	330
Qy	129	TCCTTTAGGTAGACTTTTACTACGAAAGGGGTAAAGCACATGGCTCGAAACCTTGGT	188	TCCTTTAGGTAGACTTTTACTACGAAAGGGGTAAAGCACATGGCTCGAAACCTTGGT	188
Db	331	TCCTTTGGTGGGTTTTTATTACACGAAGGTGGAACAGTAAATGGATCTCTACTCTTGT	390	TCCTTTGGTGGGTTTTTATTACACGAAGGTGGAACAGTAAATGGATCTCTACTCTTGT	390
Qy	189	TCAGCTTGTAGGGTTTCCTTTTAACCCCTTCTCTGCTATTTACTTTAAAGCCTGAGCCGTC	248	TCAGCTTGTAGGGTTTCCTTTTAACCCCTTCTCTGCTATTTACTTTAAAGCCTGAGCCGTC	248
Db	391	CCAACTTGTGGTTTTCGGATTCTCTATCTCCCTCTTTGTTTCTCTCTTCCACTC	450	CCAACTTGTGGTTTTCGGATTCTCTATCTCCCTCTTTGTTTCTCTCTTCCACTC	450
Qy	249	CAAGACTTAAACCACTTACCAAAAAAATACTCTCTTCTCTTTCGACACTATCTTTAGTGTA	308	CAAGACTTAAACCACTTACCAAAAAAATACTCTCTTCTCTTTCGACACTATCTTTAGTGTA	308
Db	451	TTCTTC-----ATCTTCTTGTTCTTTCAAGACTCTGGTTTGGATTTA	492	TTCTTC-----ATCTTCTTGTTCTTTCAAGACTCTGGTTTGGATTTA	492



QY 664 AATACACATTCAGGCTATTTTACACATGGCCACATATCCGTCTATGGTAGCTACTTGT 723  
Db 844 AGGAGACCTTCTCAGTGTGTGTAACATGCAGATATATACAGCTCTCGTGGCAACATG 903  
QY 724 TAGTGTGTAGGAGCTTTTGGAAAGTGGTGGTGGAAAGAGCTGAGTACAGAAATGGA 783  
Db 904 GCTTCTCTGTGTGGGTATTTGCAAGTGGTGAATGATGACTTTTACAAGGAGAGATGCAT 963  
QY 784 GAGTTTCAACTAGGAAAGCTACATACATTTTGTATAAATCATCGGTTTCAACGATATCATGG 843  
Db 964 GCAITCCAGTCTGGGAAGCTGATATGTAATGACACTGCTGTGACGGCTATATCTTGG 1023  
QY 844 CAAGCTTCTGATTGGAGTGTGGTTTGTGATTATCGAAGTTTCAATCGCTTTTTCAT 903  
Db 1024 CAGATGGATGGAGTAAGATTATAGCTATGCTATGCTGATGGCCATTTGGGGATTTATGTCATAT 1083  
QY 904 GTC 906  
Db 1084 GGC 1086

## RESULT 15

ADA70748  
ID ADA70748 standard; DNA; 2175 BP.

XX AC ADA70748;

XX DT 20-NOV-2003 (first entry)

XX DE Rice gene, SEQ ID 4071.

XX KW Plant; bacterial infection; fungal infection; viral infection; rice;  
XX gene; ds.

OS Oryza sativa.

XX PN WO200300898-A1.

XX PD 03-JAN-2003.

XX PF 22-JUN-2001; 2001WO-IB001105.

XX PR 22-JUN-2001; 2001WO-IB001105.

XX PA (SYGN ) SYNGENTA PARTICIPATIONS AG.

XX PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;  
XX Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;

XX DR WPI; 2003-175290/17.

XX PT Identifying at least one gene involved in plant resistance or response to  
XX pathogenic infection for conferring resistance or tolerance to a plant to  
XX bacterial, fungal or viral infection by determining or detecting plant  
XX gene expression.

XX PS Claim 6; SEQ ID NO 4071; 899pp; English.

XX The present invention relates to a method (M1) for identifying genes  
XX involved in plant resistance or response to pathogenic infection. M1  
XX comprises identifying a gene whose expression is significantly altered in  
XX the incompatible interaction of plant gene expression relative to  
XX expression of the gene in an uninfected plant, in a mutant plant that  
XX does not express a gene associated with response to pathogenic infection,  
XX or in a corresponding incompatible or compatible interaction. (M1) is  
XX useful for conferring resistance to resistance or tolerance to a plant to  
XX bacterial, fungal or viral infection. The present sequence was used to  
XX illustrate the invention.

XX SQ Sequence 2175 BP; 363 A; 701 C; 671 G; 439 T; 0 U; 1 Other;

Query Match 14.2%; Score 162.2; DB 8; Length 2175;

Best Local Similarity 50.8%; Pred. No. 1.7e-33;  
Matches 460; Conservative 0; Mismatches 418; Indels 27; Gaps 2;

QY 70 TGGAGGTTAAGAGTGTCTCTATGTCACTCTCTCTAGCTGGAGACAAATAGCCACT 129  
Db 136 TGGTGGCGATGGTGGCAGTGGAGCTCTTCTTCTCATCGCGGTGACATCGCGGACA 195  
QY 130 CTCTTAGGTAGACTTTTACTAGAAAAGGGGTAAGGACATGGCTCGAAACCTTGTT 189  
Db 196 CTGTGGGCGAGATACTACTACACCAAGGCGCGCAGCAAGTGATATCGGCGTTCTGTG 255  
QY 190 CAGCTTGTAGGGTTTCTTTAAACCTTCTTGTCTATTATTACTTTAAAGCCTGAGCCGTC 249  
Db 256 CGGAGCGCGGCTTCCGATACTGTCTTCAACCTCTTCTTCTTCCCTCGAAGTCAACG 315  
QY 250 AAGACTAAACCAATTAACAAAACAACTACTTCTTCTTGTGACACTATCTTTAGTGTAT 309  
Db 316 TCCTCTCTGCACCAACA-----CTCCCATGGCCAGCTCGCGTGATATAC 360  
QY 310 ATTGGACTTGGCTTGTGTGGACATGTATTTTGTACTCATTTGGGCTACTTTAC 369  
Db 361 ATCGTCTGGGCTCATCATCGCGCGCAGACATGATGTACTGTGGGCTCAAGTAC 420  
QY 370 CTTCTGTCTCAACTTTCTTCTTGTATCTCTCGTTCGCAATTGGCTTTTAAACGCGCTCTTC 429  
Db 421 CTCCCGCTTGCACCTACTCGTCACTGTGCCAGCAGCTCGGTTCAATGTCTCTTC 480  
QY 430 TCTTACTTCTTAAACTCACAAAAAATCACACCAATTTATATCAATTCACATTTGTTCTTTA 489  
Db 481 TCATACGTGTCAACTCCCAAGAGTCACTCTCTGATATTTCAACTCCGTCGTCTGCTC 540  
QY 490 ACCATATCTTACACTTCTTGTATTCACATGAACCAAGATCTCCCTCTTCTACTTCA 549  
Db 541 ACCATGTCCGCTTGGCTCATCGGAGTCAGCAAGAGTCTCAGGGGGTCAACG----- 592  
QY 550 AAGTCCGACGCAAGTCAAGTATGTATTGGATACATCTGCGCGGTGGTAGTCTCAGCT 609  
Db 593 ----GCGTCTCGGAGGGAAGTATCTGCTCGGTTTCTGTGCTGAGCTGGGGGCGTCTGC 648  
QY 610 GGTATTCTCTGGTGTCTTTCTTTAAACAGATTACCGGTTTCAAAAGATTCTTAAAGAAATAC 669  
Db 649 ACCTACTCGTGATCTCTCGCGCTGATGCAGCTCACTTCGAGACCATCATCAAGAGCAC 708  
QY 670 ACATTCAAGGCTATTTTAGACATGGCCACATATCCGTCTATGTGTAGTCTTGTGTAGTT 729  
Db 709 ACCTTCTCAGCGCTCTCTCAACATGCAAGTCTACACGCGCTCGTGGCGACGGCGCGCTCG 768  
QY 730 GTGTAGGACTTTTGGAAAGTGGTGGTGGAAAAAGCTGAGTACAGAAAAATGGAAGATTT 789  
Db 769 GTGTGCGGCTGTTCGCGAGCGCGAGTGGAGGTGCTGAGGGGGAGAGATGAAACGCTTC 828  
QY 790 CAACTAGGAAAAAGCTCATATACATTTTGAATAAACATCGGTTCAAGATATCATGGCAAGCT 849  
Db 829 AGGTGGGGCAGATTCTCTTACCTGATACGCTGTGTGGGCGCGCTGTCTGGCAGGTG 888  
QY 850 TGTTTGAATGGAAGTGTGGTTTGTATTCGAAGTTTTCATCGCTTTTTCATGTGTCATA 909  
Db 889 GCCAACATCGGGGTGCTCGGCGCTCATCTTCAGGTGTGCGGCGCTTCTTCTCAAGCTGATC 948  
QY 910 AGCACTCTTTTGTATCCAGTGTGCGCTGTCTTGTCTTCTTCTTCTTCCGTGATGAGATG 969  
Db 949 AGCACGGTGTCACTGCGGCTATCCCGTCTTCTCGGCTGTGGTGTTCACAGACAGATG 1008  
QY 970 AGTGG 974  
Db 1009 AACGG 1013

Search completed: November 1, 2004, 07:19:49  
Job time : 608.255 secs

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	1145	100.0	1145	6	AX033546	AX033546 Sequence
2	1145	100.0	58427	8	ATF15J5	AL110123 Arabidops
3	1145	100.0	82697	8	ATF9A21	AL021713 Arabidops
4	1145	100.0	194143	8	ATCHRIV48	AL161548 Arabidops
5	376.6	32.9	73179	8	AP006383	AP006383 Lotus cor
6	320.6	28.0	1106	8	AF370622	AF370622 Arabidops
7	318.4	27.8	1165	8	AY096558	AY096558 Arabidops
8	318.4	27.8	1428	8	AY074546	AY074546 Arabidops
9	318.4	27.8	3387	6	AX507517	AX507517 Sequence
10	318.4	27.8	3387	6	AX652015	AX652015 Sequence
11	317.2	27.7	1194	6	AX033548	AX033548 Sequence
12	308.6	27.0	1071	6	AX033550	AX033550 Sequence
13	304.8	26.6	1251	8	BT003358	BT003358 Arabidops
14	304.8	26.6	1372	8	AY062782	AY062782 Arabidops
15	304.8	26.6	115421	8	T12C22	AC020576 Sequence
16	298.4	26.1	1047	6	AX412868	AX412868 Sequence
17	298.4	26.1	1081	6	AX033549	AX033549 Sequence
18	294.4	25.7	1293	6	AX033547	AX033547 Sequence
19	281.8	24.6	1476	8	AK071254	AK071254 Oryza sat

Db	121	ATAGCCACTCTCTTAGGTAGACTTTACTACGAAAAGAGCGGTAAAGACACATGGTGC	180
Ov	181	ACCTTCGCTCAGCTGCTAGGGCTTTCCTTTTACCCCTTCCTTCGCTATATATATTAAGCCT	240



```

/number=3
complement (4116. .4281)
/gene="F15J5.10"
/number=4
complement (4282. .4724)
/gene="F15J5.10"
/number=4
complement (4725. .5023)
/gene="F15J5.10"
/number=5
6767. .6789
9077. .13984
/gene="F15J5.20"
join(9077. .9301,9805. .9859,9938. .10113,10198. .10419,
10487. .10725,10806. .11331,11419. .11692,11793. .12736,
12827. .13093,13169. .13432,13508. .13984)
/gene="F15J5.20"
/notes="similarity to multidrug resistance protein 1,
Aspergillus fumigatus, U62933
Contains ABC transporters family signature,
ABC Transporter [MSGGQKQRIAIARAI][LSGGQKQRIAIARAI],
ATP/GTP-binding site motif A (P-loop), Atp_Gtp_A
[GGSGSGKS][GGSGSGKS]"
/codon_start=1
/product="multidrug resistance
protein/P-glycoprotein-like"
/protein_id="CAB53646.1"
/db_xref="GI:5816991"
/db_xref="GOA:Q98VW6"
/db_xref="TrEMBL:Q98VW6"
/translation="MEEKSSKNDGNGKVSFFKLFSFADKTDVVLMVTGTTAAAGNG
LTQPMFLIFGOLINAFGTDPDHVREVMKVFVILAVYSCVVAFLQVSCWMTG
ERQSGATIRLKLRLQDIFYDETNTGIVGRMSGDTILIODAMGEKVKFTOLL
CTPLGGFAFAFKGPLLACVLCSCPLIVIAAGASLIMSKMAGRGVAVAEAGNVVE
QTVGAIRTVVATGEKQATEKYESKLEIAYKTWQGLISGFGCLTMLAVIFCSGLA
VYGAKLIMEKGVNGQVINIIPAVLTGMSGQSPSLNFAAGRAAFKFFETIKR
SPKIDAYMSGVLEDIGDIELKDVIFRYPARPDVQIPAGFSLFVPMGKTVALVGSQ
GGKSTVLSIERFYDPESGQVLIDNIDLKQLKWRISKIGLVQEPVLFATTIKEN
IAYGKEDATQBITAIELAENAAKFDIKLPQGLDWTMGHGTQMSGGQKQRIAIARAI
LNPKILLDEATSDALDAESRIVQDALVNLMSNRTVVVAHRLTTIRTADVIAVHQ
GKIVEKTHDEMIDQPEGAYSQVLRQEGSKEATESERPETSLOVERSGSRLSSAM
RRSVNRSSSRHSFSLASNMFPGVNVNQTDMEDEENNVRHKVSLKRLAHLNKE
IPVLVLSIAAMVHGTVFPIFGLLSSSINMFYEPAKILKDSHFALYIYALGUTNF
VMIPNYPFPGIAGGKLIKIRISMCFKVHVEIISWFDOTANSRYNFIYIINRILY
VLILFICVLLPPEVRLERECSDASTVSLVGDALALIVQNIATVTTGLIIAFTANWI
LALIVALSPEFTIVIQGYAOTKLTGFSADAKMYEASOVANDAVSSIRTVASFCAEE
KWDMLYQQKCDGPKNGVRLGLLSGAGFGFSFFFLYCINCVCVFSAGLIQIKATFG
EVFKVFALTUWAIGVQSQTSAWAPDSNKAQSAISIFDLDSTPKIDSSSDGTTLQN
VNGDIEFRVSRYPMPRPDVQIFRDLCLTIPSGKTVALVGSQSGKSTVISMIRFYN
PDGKILIDQVEITQIFKLSWLQOMGLVQSEPILFNETIRSNIAVYKGTGGATEEIIA
AAKAAAHNFISLPGQYDTSVGERGVQLSGGQKQRIAIARAILKDPKILLDEATSA
LDAESRVVDALDRVXG"
9077. .9301
/gene="F15J5.20"
/number=1
9302. .9804
/gene="F15J5.20"
/number=1
9805. .9859
/gene="F15J5.20"
/number=2
9860. .9937
/gene="F15J5.20"
/number=2
9938. .10113
/gene="F15J5.20"
/number=3
10114. .10197
/gene="F15J5.20"
/number=3
10198. .10419
/gene="F15J5.20"
/number=4

```

```

intron
/gene="F15J5.20"
/number=4
10487. .10725
/gene="F15J5.20"
/number=5
10726. .10805
/gene="F15J5.20"
/number=5
10806. .11331
/gene="F15J5.20"
/number=6
11332. .11418
/gene="F15J5.20"
/number=6
11419. .11692
/gene="F15J5.20"
/number=7
11693. .11792
/gene="F15J5.20"
/number=7
11793. .12736
/gene="F15J5.20"
/number=8
12737. .12826
/gene="F15J5.20"
/number=8
12827. .13093
/gene="F15J5.20"
/number=9
13094. .13168
/gene="F15J5.20"
/number=9
13169. .13432
/gene="F15J5.20"
/number=10
13433. .13507
/gene="F15J5.20"
/number=10
13508. .13984
/gene="F15J5.20"
/number=11
14541. .16535
/gene="F15J5.30"
complement (join(14541. .14660,14757. .14831,14919. .15008,
15217. .15414,15497. .15598,15696. .15850,15955. .16024,
16148. .16270,16476. .16535))
/gene="F15J5.30"
complement (join(14541. .14660,14757. .14831,14919. .15008,
15217. .15414,15497. .15598,15696. .15850,15955. .16024,
16148. .16270,16476. .16535))
CDS
Query Match 100.0%; Score 1145; DB 8; Length 58427;
Best Local Similarity 100.0%; Pred. No. 2.3e-267;
Matches 1145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CTGTTTTCATGTGTTGATGGTAGAACCTCGAAGGGAATTTTCAACAGAAGAGAGAAGTCCAC 60
DB 54933 CTGTTTTCATGTGTTGATGGTAGAACCTCGAAGGGAATTTTCAACAGAAGAGAGAAGTCCAC 54992
QY 61 AAGTACTCTTGGAGGTTAAGAGTGTCTCTATGTGCTACTCTCTCTTGTAGCTGGAGAGACA 120
DB 54993 AAGTACTCTTGGAGGTTAAGAGTGTCTCTATGTGCTACTCTCTCTTGTAGCTGGAGAGACA 55052
QY 121 ATAGCCACTCTCTTAGGTAGACTTTTACTACGAAAAGCGGTAAAGACACATGGCTCGAA 180
DB 55053 ATAGCCACTCTCTTAGGTAGACTTTTACTACGAAAAGCGGTAAAGACACATGGCTCGAA 55112
QY 181 ACCTTGGTTACGCTTGTAGGGTTTCCTTTAACCCCTTCCTTCTATTATTACTTAAAGCCT 240
DB 55113 ACCTTGGTTACGCTTGTAGGGTTTCCTTTAACCCCTTCCTTCTATTATTACTTAAAGCCT 55172
QY 241 GAGCGGTCGAAGACTAAACCACTTACCAAAAAAACAATACTTCTTCTTCTTGACACTATCT 300

```





```
/db_xref="GOA:O49721"
/db_xref="TrEMBL:O49721"
/translation="AFVAAMDKACKSSSSVNLIIPRGFSVGLRFGPCTNVSMLT
RYKASTDLKSPSGGWQFGWINGLITLGGTDFGQALWPNPNNCTSDNCKLLP
TSLKFFVMRTVVRIRLSSVNSKFFPHIALVECRDFKGTRLNITAPSDPNTDGIHTRS
SNYIFRSHIATGDDCVSGQNSQITITISKCGPHGHSVGLSPNPKDGVNGLV
KDKISGTINGIRIKTIANSPGLSAATMTFENIMNVNTPIIIDQSCPFSSCISN
VPKNSIEIYFNRIKRTSSSLVAVQLHCSRGMPCCKVYLENVHLDLSSSDGGRKQSS
NRGNEAVSSSCRNVRYNGTQIPPPCH"
exon      4778..4921
           /gene="T9A21.20"
           /number=1
intron    4922..5022
           /gene="T9A21.20"
           /number=1
exon      5023..5187
           /gene="T9A21.20"
           /number=2
intron    5188..5262
           /gene="T9A21.20"
           /number=2
exon      5263..5552
           /gene="T9A21.20"
           /number=3
intron    5553..5637
           /gene="T9A21.20"
           /number=3
exon      5638..5866
           /gene="T9A21.20"
           /number=4
intron    5867..5950
           /gene="T9A21.20"
           /number=4
exon      5951..6211
           /gene="T9A21.20"
           /number=5
gene      6969..8087
CDS       join(6969..7148,7191..8087)
           /gene="T9A21.30"
           /note="similarity to W30DMY32F, W25DMY32F"
           /codon_start=1
           /product="putative protein"
           /protein_id="CAA16790.1"
           /db_xref="GI:2832692"
           /db_xref="TrEMBL:O49722"
           /translation="MYEPGKFEERSHKYWRLRVSLVYVLLLAGETIALLGLR
YKGGSTWLETUQPEPSKTIITKKTSFTLSLUSLVYIGLLVAGHCLYSPGL
LYLPVFTSLISQAFNAVFSYFLNSQKITPFIINSLVLTISLTLLVIOHEPSP
SSTSKSAASKYVIGYICAVGSAGYSLVSLTDYAFKILKKTPTKALIDMATYPSM
VATCVVVGKFGSGGWKLSLSTEMEERFOLGKSYILINIGSTISWQACLIGSYNLIIEV
SSLFSNVI STLCLPVVPLAVVFRDEMSGIKLVAMFLAINGFVSYGHYVNDRKE
EDQLPQSKEEERQKQVDIHQA"
exon      6969..7148
           /gene="T9A21.30"
           /number=1
intron    7149..7190
           /gene="T9A21.30"
           /number=1
exon      7191..8087
           /gene="T9A21.30"
           /number=2
gene      8612..13959
CDS       join(8612..8838,8884..9860,10782..11155,11210..11915,
           12857..13959)
           /gene="T9A21.50"
           /note="similarity to W30DMY32F, W15DMY32F
           contains EST gb.T42237"
           /codon_start=1
           /product="putative protein"
           /protein_id="CAA16792.1"
           /db_xref="GI:2832694"
```

```
/db_xref="TrEMBL:O49724"
/translation="MEITQVIYVNGKQDASRSVDYLILPANLLFLMFSDHNNIEANLT
GOENWNTMEIESSVPOSKNYKKWLRSIYVYVYENGKSTWMTGLVQLIGPVLFL
PEFSSQTKPKTEADPRKESFTILGSVYIVTGLLSANSYMSVGLLYLVSTFSL
ILASQATAPFSEFLNSOKFTFPFVNSLFLTLTISALLVNTSDENTAKVRKTVI
GLICTASAGIGLLLSVLQILRLKVKQTFSTVTDLVAYOSLVASCVLILGLFASG
EWTITSEMENYKLGKVPYMTLASIAISWQYITIGVGLIFESSVFNSTAVGLP
IYVAVAVTFHDKMNASKIFSILAIWGFISFYQHYDEKKLKTSHTSFVGDPHLLP
ABEGHTNHSQNLNLIIDHEVWTESSASVPOTENYKRWLRVSIYVIYVLCQPLA
TLRLYHENGNSYVVTLLQLIGPVLVLPFRPSRIROPKSTDNTNFSQSPFTTLA
SYLCTGLLSAYAYLSALATFAPFSYFLNSOKFTPLIVSSLLLTIVSALLVWNTDS
ENSTNSRVQYVIGFICTIGASAGIGLLLSLQMLFRKVTFKHTTSVTDLAIYQSLV
ASCVVLI GLFASGEWETLPSEMRNYKLGKSVYLTLASAASWQYITIGVGLIFESS
SVFNSITAVGLPIYVAVAVTFHDKMDSKIFSILAIWGFISFYQHYDEKKLKTSH
SHTSAVGDHLHPVERGHTNIQSDQNLNLIIDHEETSPFQTKNCKRWLRVSIYAI
FVIFCQPLATVLRLYYENGKSTVVTLLQLIGPVLVLPFRPSRIROPKSTDNTNFS
QSPFTTLASVYLTGLLSAYAYLSAVGLLILPVSTFSLILASQATAPFSEFLNS
QKFTPLIVNSLFLTIVSALLVWNTSDNTNFSQSPFTTLASVYLTGLLSAYAYLS
IQLFRKVTFKHTTSVTDLAIYQSLVLTLASAASWQYITIGVGLIFESSVFNSTAV
YLTLASAASWQYITIGVGLIFESSVFNSTAVGLPIYVAVAVTFHDKMDSKIFSIL
IPSIILAIWGLFSFYQHYDEKKLKTQTKVEBETQL"
```

exon	8612..8838	/gene="T9A21.50"	/number=1
intron	8839..8883	/gene="T9A21.50"	/number=1
exon	8884..9860	/gene="T9A21.50"	/number=2
intron	9861..10781	/gene="T9A21.50"	/number=2
exon	10782..11155	/gene="T9A21.50"	/number=3
intron	11156..11209	/gene="T9A21.50"	/number=3
exon	11210..11915	/gene="T9A21.50"	/number=4
intron	11916..12856	/gene="T9A21.50"	/number=4
exon	12857..13959	/gene="T9A21.50"	

Query Match 100.0%; Score 1145; DB 8; Length 82697;  
Best Local Similarity 100.0%; Pred. No. 2.2e-267;  
Matches 1145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	CTGTTTCATGTTGATGGTAGAACCTGAAGGGAATTTTCAACAGAGAGAGAGTCCAC	60
DB	6954	CTGTTTCATGTTGATGGTAGAACCTGAAGGGAATTTTCAACAGAGAGAGTCCAC	7013
QY	61	AAGTACTCTTGGAGGTTAAGAGTGCTCTCTATGTACACTCTCTCTTACTCTGAGAGACA	120
DB	7014	AAGTACTCTTGGAGGTTAAGAGTGCTCTCTATGTACACTCTCTCTTACTCTGAGAGACA	7073
QY	121	ATAGCCACTCTCTTAGGTAGACTTTTACTACGAAAAAGCGGTAAAGACACATGGCTCGAA	180
DB	7074	ATAGCCACTCTCTTAGGTAGACTTTTACTACGAAAAAGCGGTAAAGACACATGGCTCGAA	7133
QY	181	ACCTTGGTTACGCTTGTAGGGTTTCCTTTAACCCCTCTCTGCTATATATTACTTAAGCCT	240
DB	7134	ACCTTGGTTACGCTTGTAGGGTTTCCTTTAACCCCTCTCTGCTATATATTACTTAAGCCT	7193
QY	241	GAGCCGTCCAGACTAAACACCATTTACCAAAAAAACTACTTCTCTCTTGTGACACTATCT	300
DB	7194	GAGCCGTCCAGACTAAACACCATTTACCAAAAAAACTACTTCTCTCTTGTGACACTATCT	7253
QY	301	TTAGTGTATATTGGACTTGGCTTGCTTGTGTGGACATTGTATTTTGTACTCATTTGGG	360

Db	7254	TTAGTGATATATTGGACTTGGCTTGCTTGTTGCTGGACATGTATTATTGTACTCAATTGGG	7313
QY	361	CTACTTTTACCTTCCTGCTCAACTTTCTCTTTGATCTCTGCGTCGCAATTGGCTTTTAAC	420
Db	7314	CTACTTTTACCTTCCTGCTCAACTTTCTCTTTGATCTCTGCGTCGCAATTGGCTTTTAAC	7373
QY	421	GCGCTCTCTTCTTACTTCTTAACTCAACAAAAACACACCACTTTATATCTCAATTCACATT	480
Db	7374	GCGCTCTCTTCTTACTTCTTAACTCAACAAAAACACACCACTTTATATCTCAATTCACATT	7433
QY	481	GTTCTCTTAAACCATATCTTCTACACTTCTTGTTATCCAAATGAACCAAGTTCCTCCTCT	540
Db	7434	GTTCTCTTAAACCATATCTTCTACACTTCTTGTTATCCAAATGAACCAAGTTCCTCCTCT	7493
QY	541	TCCTACTTCAAGTCGCGCAAGTCCCAAGTATGTGATGGATACATCTGCGCGTCCGT	600
Db	7494	TCTACTTCAAGTCGCGCAAGTCCCAAGTATGTGATGGATACATCTGCGCGTCCGT	7553
QY	601	AGCTCAGCTGGTATTCTCTGCTGCTTTCTTTAAACAGATTACGCCGTTGAAAAGATTCTA	660
Db	7554	AGCTCAGCTGGTATTCTCTGCTGCTTTCTTTAAACAGATTACGCCGTTGAAAAGATTCTA	7613
QY	661	AGAAATACACATTCAGGCTATTTTAGACATGGCCACATATCCGTCATGTAGTACTACT	720
Db	7614	AGAAATACACATTCAGGCTATTTTAGACATGGCCACATATCCGTCATGTAGTACTACT	7673
QY	721	TCGTAGTGTGCTAGGACTTTTTCGAAGTGTGCTGGTGAAGAACTCAGTACAGAAATG	780
Db	7674	TCGTAGTGTGCTAGGACTTTTTCGAAGTGTGCTGGTGAAGAACTCAGTACAGAAATG	7733
QY	781	GAAGAGTTTCAACTAGGGAAGCTCATACATTTTGTATAAATCATCGGTTCAACGATATCA	840
Db	7734	GAAGAGTTTCAACTAGGGAAGCTCATACATTTTGTATAAATCATCGGTTCAACGATATCA	7793
QY	841	TGGCAAGCTTGTGATTTGGAAGTGTGGTTGATATCGAAGTTTCATCGCTTTTTCCTCC	900
Db	7794	TGGCAAGCTTGTGATTTGGAAGTGTGGTTGATATCGAAGTTTCATCGCTTTTTCCTCC	7853
QY	901	AATGTCATAAGCACCTTTGTTTACAGTGTGCGCTGTTCTTGCGTGTGCTCTCTCCGT	960
Db	7854	AATGTCATAAGCACCTTTGTTTACAGTGTGCGCTGTTCTTGCGTGTGCTCTCTCCGT	7913
QY	961	GATGAGTCAAGTGAATCAAGTTGTTGCAATGTTTTTGGCATCTGGGGGATTTGTTTCT	1020
Db	7914	GATGAGTCAAGTGAATCAAGTTGTTGCAATGTTTTTGGCATCTGGGGGATTTGTTTCT	7973
QY	1021	TATGTTTATCAGCATTTATGTCATGATAGAAGCCAGAGAAGCAAGAGCTTCCTCAG	1080
Db	7974	TATGTTTATCAGCATTTATGTCATGATAGAAGCCAGAGAAGCAAGAGCTTCCTCAG	8033
QY	1081	TCTAAGAAGAGAAGAAACAAACAAAGTAGATACCATTTGTTCCAAAGCTTAGGCAAG	1140
Db	8034	TCTAAGAAGAGAAGAAACAAACAAAGTAGATACCATTTGTTCCAAAGCTTAGGCAAG	8093
QY	1141	ATCCA 1145	
Db	8094	ATCCA 8098	
RESULT 4			
ATCHRIV48			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
Arabidopsis thaliana (thale cress)			
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;			
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.			
REFERENCE			
1 (bases 1 to 111084)			

AUTHORS	Murphy,G., Ridley,P., Hudson,S., Mewes,H.W., Lemcke,K. and Mayer,K.F.X.
JOURNAL	Unpublished
REFERENCE	2 (bases 107966 to 194143)
AUTHORS	Hilbert,H., Braun,M., Holzer,B., Brandt,A., Duesterhoeft,A., Mewes,H.W., Lemcke,K. and Mayer,K.F.X.
JOURNAL	Unpublished
REFERENCE	3 (bases 1 to 194143)
AUTHORS	EU Arabidopsis sequencing, project.
TITLE	Direct Submission
JOURNAL	Submitted (10-MAR-2000) MIPS, at the Max-Planck-Institut fuer Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail: lemcke@mips.biochem.mpg.de, mayer@mips.biochem.mpg.de, project
AUTHORS	Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK, E-mail: michael.bevan@bbsrc.ac.uk
COMMENT	Information on performance of analysis and a more detailed annotation of this entry and other sequences of chromosomes 3, 4 and 5 can be viewed at: http://www.mips.biochem.mpg.de/proj/thal/ this fragment has an overlap with ATCHRIV47 at the 5' end and an overlap with ATCHRIV49 at the 3' end.
FEATURES	Location/Qualifiers
source	1..194143
	/organism="Arabidopsis thaliana"
	/molecule="genomic DNA"
	/variety="Columbia"
	/db_xref="taxon:3702"
	/chromosome="4"
	6377..9293
gene	/gene="AT4g18120"
CDS	join(6377..6796,7122..7238,7347..7501,7578..8456,8543..8627,8716..8784,8871..9293)
	/note="AT4g18120"
	/note="similarity to Mei2-like protein, Arabidopsis thaliana, gb:D86122
	contains EST gb:W43784, AT998063.1"
	/codon_start=1
	/product="putative protein"
	/protein_id="CAB78814.1"
	/db_xref="GI:7268605"
	/db_xref="GOA:Q9SVV9"
	/translation="MASGLDHFSGSIGNMLDDGDSHPIGNMLPDBEELFSGMLDDLNLSSLPATLDLEDYDLFGSGGLELETDYPDYSLNKSPRMGFADSNVDNVPQNIQFQVNSIAGEHPYGEHPSTLFRVINSNVEDSEIQALPEDNPKVDNQGLVFNFLAPSVNRDLNIFGVYGEIKQIETPNKGRHKHVEFEDVRSADAALKALNTEAGKILKHSRPGGARRNMQLMQLNVEPLASPIGNMNRNSPIDRPLQSF SKSPIFNLSPTKNIRYPEFSMTASVYNDQGRFSLDHLDFSSSYNNASHKASTFQQPSFGVSFSGLSNHPHSHVETLSGSEFLWGSPPSSAWPVNPFSSNRNHRFPYSAQNSGLHOLHHIGSAPGFFPRSPETSMGSVAFRGASGMNAQRNLRSTSPNFKMLSAPRRSOLFTGNGSYLWPAATVMSIDDDPLEDSNQFDNSGNQADIKIQFOLDLSKIMRGEPRPTLMIKNIPKYNRMWLLAAIDENKSGTYDELYLPIDFNKCNVGYAFINMVS PKTIALIYFAFNKKWDKFNSEKVASLAYARIQKALIAHFQNSSLMNEDRRCCQPIVFDGSESKYPPVNNLLQQTHDLKMIYILFCVYLMNWLTHCLTFETRSRSGEFAARASNSGTFLIFLSARHIKADSQNVHVIASGLKRYQFLG"
exon	6377..6796
	/gene="AT4g18120"
	/number=1
intron	6797..7121
	/gene="AT4g18120"
	/number=1
exon	7122..7238
	/gene="AT4g18120"
	/number=2
intron	7239..7346
	/gene="AT4g18120"
	/number=2
exon	7347..7501
	/gene="AT4g18120"
	/number=3
intron	7502..7577
	/gene="AT4g18120"
	/number=3

```
exon      7578. .8456
           /gene="AT4g18120"
           /number=4
intron    8457. .8542
           /gene="AT4g18120"
           /number=4
exon      8543. .8627
           /gene="AT4g18120"
           /number=5
intron    8628. .8715
           /gene="AT4g18120"
           /number=5
exon      8716. .8784
           /gene="AT4g18120"
           /number=6
intron    8785. .8870
           /gene="AT4g18120"
           /number=6
exon      8871. .9293
           /gene="AT4g18120"
           /number=7
gene       9594. .13230
           /gene="AT4g18130"
gene       complement (join(9594. .9837,9944. .10237,10323. .11130,
           11238. .13230))
           /gene="AT4g18130"
CDS        complement (join(9594. .9837,9944. .10237,10323. .11130,
           11238. .13230))
           /gene="AT4g18130"
           /note="Contains ATP/GTP-binding site motif A (P-loop)
AA634-641;Phytochrome chromophore attachment site
AA317-326
           contains EST gb:T21020, AA395462"
           /codon_start=1
           /product="phytochrome E"
           /protein_id="CAB78815.1"
           /db_xref="GI:7268606"
           /db_xref="GOA:P42498"
           /db_xref="Swiss-Prot:P42498"
           /translation="MGFSSSSAASNMKPPQKSNATQYSDVDAALFADPAOSIVTGKS
FNYKSVISPNHVPDEHTAYLSNIQRGLVQPGCLIAVEEPSFRILGSDNSSDF
LGLLSPTSGSEDFKVLGIGIDARTLFTPSSGASLKAASFTEISLNPVLVHSR
TTQKPFYAILHRIDAGIVMDLEPAKSGDPALTLAGAVQSLAVRAISRLOSPLGDI
GALCDTVQVQLTGDRVMVYQFHEDDHGEVSEIRSDLEPYLGHYPATDIPQA
ARFLPKQNRVLCNATPPKVVQSEELKRPCLVNSTFLRAPHCHGYMANMGSA
SLALAIVVKGKDSKLGLVVGHCSPRYVPPPLAYACEFLMQAFLQLOLQASQ
LAEKAMRTQTLCDMLARDTVSAIVTQSPGIMDLVKGCDGALYYKGCWLGVPTPE
SQVDLVNVLVNHGDDSTGLTTLTSLVDAGYPGALSGLDVGVAABFSSKDYLLWF
RSNTASAIKMGAKHHPKDDKODAGMRPSSFTAFLEVAKRSLEPWEISEIDAIHSR
LIMRESFTSSRPVLSGNGVARDANELTSFVCEMVRVITATAPIFGVDSGCGINGWK
KTAEMTGILLASAMGKSLADEIVQESRAALESLLKALOGEEESVMLKLRKFGQNN
HPDYSDDVCLVNSTSRDYTENIIGVCFVGODITSEKAITDRFIRLOGDYKTIQSL
NPLIPIFASDENACCSSEWNAAMEKLGTGSKHEVIGKMLPGEVFGVCKVKQDSLTK
PLISLYQGIAGNVPESSLVFEFNKEGKYIEASLTANKSTNIEGVIRCFPLQIINK
EGLSCPSELKSAQSLNELTVYRQIKNPLNGIRFAHKLSEISASQRFLETSDA
CEKQITTIESTDLKISIEBKLOLETEEFRELENIIDTIIISOVMIILRNRQSLRVEVA
EBIKTLPLNGDRVKLQILADILRNIVNHAPPNPSWVGISISPGQELSRDNGRYTHLO
FEMIHPGKGLPSEMLSDMEFTDGNVTPDGLGLKLSRLLEQMGRGVSYVREDERCF
QVDLQKTMVGESRGTGSSSIK"
           complement (9594. .9837)
           /gene="AT4g18130"
           /number=1
intron    complement (9838. .9943)
           /gene="AT4g18130"
exon      complement (9944. .10237)
           /gene="AT4g18130"
           /number=2
intron    complement (10238. .10322)
           /gene="AT4g18130"
           /number=2
exon      complement (10323. .11130)
           /gene="AT4g18130"

/number=3
complement (11131. .11237)
/gene="AT4g18130"
/number=3
complement (11238. .13230)
/gene="AT4g18130"
/number=4
/gene="AT4g18140"
complement (join(14593. .14751,14921. .15070,15338. .15639,
15915. .16055,16141. .16312))
/gene="AT4g18140"
complement (join(14593. .14751,14921. .15070,15338. .15639,
15915. .16055,16141. .16312))
/gene="AT4g18140"
/note="similarity to HYA22 protein, Homo sapiens,
PIR2:JC5707
           contains EST gb:R65235"
           /codon_start=1
           /product="putative protein"
           /protein_id="CAB78816.1"
           /db_xref="GI:7268607"
           /db_xref="TrEMBL:Q9SVV8"
           /translation="MKSKISKDSLRERKRVLGVCXKPKHVDXNCSFDPKISRMEILLVT
ATQNCNNEEQNIIGLSDSTRDVTVNEPQDLIVEDSSKNASNMDTIFSPVLDDLYSET
DREPSVLLDAAEQYMLLPYLEDRSASSDDVKSDENRINRSRSLKNLEAENHTEAEQT
EDFPQIFLRNQPELADVFNYPVDMQOPRDSPKRKAVTFVLINMLLDILDPDGKFSV
QRFRDSCILSDGIYTKDLTVLGLDLAKVAIVDNCQPVYRLQINNGIFIKSWYDDPTD
DGLITLIPFLLETADANDVRPIAKRFGNKE"
           complement (14593. .14751)
           /gene="AT4g18140"
           /number=1
           complement (14752. .14920)
           /gene="AT4g18140"
           /number=1
           complement (14921. .15070)
           /gene="AT4g18140"
           /number=2
           complement (15071. .15337)
           /gene="AT4g18140"
           /number=2
           complement (15338. .15639)
           /gene="AT4g18140"

Query Match      100.0%; Score 1145; DB 8; Length 194143;
Best Local Similarity 100.0%; Pred. No. 2e-267;
Matches 1145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGTTTCATGTGTTGATGGTAGAACCTGAAGCGGAAATTTTCAACAGAGAGAGAGAGTCAAC 60
Db 35341 CTGTTTCATGTGTTGATGGTAGAACCTGAAGCGGAAATTTTCAACAGAGAGAGAGTCAAC 35400

Qy 61 AAGTACTCTTGGAGGTTAAAGAGTGCTCTCTATGTCACTCTCTCTTATAGCTTGGAGAGACA 120
Db 35401 AAGTACTCTTGGAGGTTAAAGAGTGCTCTCTATGTCACTCTCTCTTATAGCTTGGAGAGACA 35460

Qy 121 ATAGCCACTCTCTAGGTAGACTTTTACTACGAAAAGGCGGTAAAGACACATGGCTCGAA 180
Db 35461 ATAGCCACTCTCTAGGTAGACTTTTACTACGAAAAGGCGGTAAAGACACATGGCTCGAA 35520

Qy 181 ACCTTGGTTACAGTTGTAGGGTTTCCTTTAACCCCTTCCTTCTATATTATTACTTAAAGCCT 240
Db 35521 ACCTTGGTTACAGTTGTAGGGTTTCCTTTAACCCCTTCCTTCTATATTATTACTTAAAGCCT 35580

Qy 241 GAGCCGTCCAAGACTAAACCAATTACCAAAAAAACAATCTTCTTCTTCTTCTGACACTATCT 300
Db 35581 GAGCCGTCCAAGACTAAACCAATTACCAAAAAAACAATCTTCTTCTTCTTCTGACACTATCT 35640

Qy 301 TTAGTGATATATTGACATTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGG 360
Db 35641 TTAGTGATATATTGACATTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGG 35700

Qy 361 CTACTTTACCTTCCTGTCTCAACTTTCTTTGTATCTCTGCGTCGCAATTTGGCTTTTAAAC 420
```

35701	Db	CTACTTTACCTTCTCTGTCTCAACTTTCTCTTTGATCTCTGCGTCGCAATTGCGTTTAACT	35760
421	Qy	GGCGTCTTCTTCTTACTTTCCTAAACTCAGAAAAAATCACACCAATTTATCTCAAAATTCAGTT	480
35761	Db	GGCGTCTTCTTCTTACTTTCCTAAACTCAGAAAAAATCACACCAATTTATCTCAAAATTCAGTT	35820
481	Qy	GTTCCTTTAAACCATATCTCTACACTTCTTTGTTATCCAAATCAAGAACAGAAATCTCCCTCT	540
35821	Db	GTTCCTTTAAACCATATCTCTACACTTCTTTGTTATCCAAATCAAGAACAGAAATCTCCCTCT	35880
541	Qy	TCTACTTTCAAAGTCGCGCAGCAAGTCCAAGTATGTGATTTGGATATCATCTGCGCGGTCTCGGT	600
35881	Db	TCTACTTTCAAAGTCGCGCAGCAAGTCCAAGTATGTGATTTGGATATCATCTGCGCGGTCTCGGT	35940
601	Qy	AGCTCAGCTGGTTATCTCTGCGTCTTTCTTTAAACAGATTCGCGTTTCGAAAGAAATCTTA	660
35941	Db	AGCTCAGCTGGTTATCTCTGCGTCTTTCTTTAAACAGATTCGCGTTTCGAAAGAAATCTTA	36000
661	Qy	AAGAAATACACATTTCAAGGCTATTTTAGACATGGCCACATATCCGCTATGTGTAGCTACT	720
36001	Db	AAGAAATACACATTTCAAGGCTATTTTAGACATGGCCACATATCCGCTATGTGTAGCTACT	36060
721	Qy	TGTGTAGTTGTGTAGGACTTTTTTGAAGTGTGGGTGGAAAAAGCTGAGTACAGAAATG	780
36061	Db	TGTGTAGTTGTGTAGGACTTTTTTGAAGTGTGGGTGGAAAAAGCTGAGTACAGAAATG	36120
781	Qy	GAAGAGTTTCAACTAGGGAAAAAGCTCATATTTTGATATAAACATCGGTTCAACGATATCA	840
36121	Db	GAAGAGTTTCAACTAGGGAAAAAGCTCATATTTTGATATAAACATCGGTTCAACGATATCA	36180
841	Qy	TGGCAAGCTTTTGTGATGGAAGTGTGGTTGATTCGAAAGTTTCATCGCTTTTTTCCC	900
36181	Db	TGGCAAGCTTTTGTGATGGAAGTGTGGTTGATTCGAAAGTTTCATCGCTTTTTTCCC	36240
901	Qy	AATGTCAATAAGCACTCTTTGTTTACCAGTTGTGCGCTGTTCTTGTCTGTGTCTTCCGT	960
36241	Db	AATGTCAATAAGCACTCTTTGTTTACCAGTTGTGCGCTGTTCTTGTCTGTGTCTTCCGT	36300
961	Qy	GATGAGATGAGTGGAAATCAAGTTGGTTGCAATGTTTTTGGCCATCTGGGGATTTGTTTCT	1020
36301	Db	GATGAGATGAGTGGAAATCAAGTTGGTTGCAATGTTTTTGGCCATCTGGGGATTTGTTTCT	36360
1021	Qy	TATGGTTATCAGCATATTATGTCATATAGAAAGCCAGAAAGAACCAAGAGCTTCCCTCAG	1080
36361	Db	TATGGTTATCAGCATATTATGTCATATAGAAAGCCAGAAAGAACCAAGAGCTTCCCTCAG	36420
1081	Qy	TCTAAAGAAGAAGAAACAAAAACAAAGTAGATACCATTTCATGTCCAAAGCTTAGGCAAG	1140
36421	Db	TCTAAAGAAGAAGAAACAAAAACAAAGTAGATACCATTTCATGTCCAAAGCTTAGGCAAG	36480
1141	Qy	ATCCA 1145	
36481	Db	ATCCA 36485	

RESULT 5	
AP006383/c	
LOCUS	AP006383 73179 bp DNA linear PLN 22-JUL-2003
DEFINITION	Lotus corniculatus var. japonicus genomic DNA, chromosome 6, clone:LT05I20, TM0245, complete sequence.

AEIWORKDS	FIG.
SOURCE	Lotus corniculatus var. japonicus (Lotus japonicus)
ORGANISM	Lotus corniculatus var. japonicus
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Lotaceae; Lotus.

1. *Lotus*.

REFERENCE

ASAMIZU, E., KATO, T., SATO, S., NAKAMURA, Y., KANEKO, T. and TABATA, S.  
AUTHORS  
Structural Analysis of a *Lotus japonicus* Genome. IV. Sequence  
TITLE  
Genome. IV. Sequence

FEATURES	Features and Mapping of seventy-three TAC clones which cover the 7.5 Mb Regions of the Genome
JOURNAL REFERENCE	DNA Res. (2003) In press
AUTHORS	2. (bases 1 to 73179)
TITLE	Sato, S.
JOURNAL	Direct Submission
	Submitted (07-MAY-2003) Shusei Sato, Kazusa DNA Research Institute, Department of Plant Gene Research; 2-6-7 Kazusa-kamatari, Kisarazu, Chiba 290-0818, Japan (E-mail: ssato@kazusa.or.jp, URL: <a href="http://www.kazusa.or.jp/">http://www.kazusa.or.jp/</a> , tel: 81-438-52-3935 (ex. 2337), Fax: 81-438-52-3934)
FEATURES	Location/Qualifiers
source	1. 73179 /organism="Lotus corniculatus var. japonicus" /mol_type="genomic DNA" /variety="japonicus" /db_xref="taxon:34305" /chromosome="6" /clone="LJ105120" /clone_lib="LJT library" /note="TAC clone: TM0245-synonym: Lotus japonicus"

## ORIGIN

Query Match	32.9%	Score 376.6	DB 8	Length 73179
Best Local Similarity	61.7%	Pred. No. 8.3e-81		
Matches 620	Conservative 0	Mismatches 379	Indels 6	Gaps 1
QY	59	ACAAGTACTCTTGGAGGTAAAGAGTGCTCTCTATGTCACTCTCCTCTTAGCTGGAGAGA	118	
DB	17242	AAAGTATTACCGTTGGTTTCGATTAGCCCTCTATGCAGCACTGGTGTGGTGCCAAT	1718	
QY	119	CAATAGCCACTCTCTTAGGTAGACTTTTACAGAAAAGGCGGTAAAGCACATGGCTCG	178	
DB	17182	CAGCTGCAATACTTCTGGGAAGATTGTACTATGAAAAGAGGAGAAAACAGCAAAATGGAATG	1712	
QY	179	AAACCTTGGTTCAGCTGTAGGGTTTCCTTTAAACCCCTTCCTTGTATTTACTTAAAGC	238	
DB	17122	CAACATAGCCAACTTTGCGAGGTTTCCTTATCTGTCTCTACTCTCATCTTAGCAT	1706	
QY	239	CTGAGCCGTCCAGACTAAACCAATTACCAAAAAAACTACTTCTTCTTTTGACACTAT	298	
DB	17062	CCAAGGATCTCAACACCACAAATAGCTTTTCATCAGAATCAACCTTCGTGTTCAATGCTAG	1700	
QY	299	CTTTAGTGTATATTGGACTTTGGCTTGGTGTGCTGGACATTTGATTTTGTACTCATTTG	358	
DB	17002	CATTGTCTATGTCCTCCCTTGGCCCTCTCTGGCAATTAGATTGTTATTTATCTCTCTTG	1694	
QY	359	GGCTACTTTTACTTCCCTGTCTCAACTTTCTTTTGATCTCTGGGTGCCAATTTGGCTTTTA	418	
DB	16942	GACTTTTATACCTCCCTGTCTTACTTTACACATCTATGGCTCATCTCAATTGGCTTCA	1688	
QY	419	ACGCCGTCTTCTTTACTTCTCTAACTCAACAAAATCAACACATTTATCTCAATTCAAC	478	
DB	16882	ATGCTCTCTTCTCTCTTCTCAATTCACCTGAGAAATCACACCTCTCATCACTCACTC	1682	
QY	479	TTGTTCTTTAAACATATCTTACACCTCTTGTATTCCAACATGAACCAAGATCTCCCT	538	
DB	16822	TAGTCTCTTCAACCAATTTCTCTCTCCCTCTCTGTGTTTCAAAA-----GGACTCATAT	1676	
QY	539	CTTCTACTTCAAAAGTCGCGACCCAAGTCCAAGTATGTATTGGATACATCTCGCGCGTGC	598	
DB	16768	CTGATGATTCGCGCAAGTCTCCAACAAAATATATGTATTGGATTCATATGCACAGTTG	1670	
QY	599	GTAGCTCAGCTGGTGTATTCTCTGGTGTCTTTTAAACAGATTACGCGTTGAAAAGATTTC	658	
DB	16708	CTGCATCAGCAGGCTATGGATTGGTGTCTTTCCCTTACCCTAAGTTGCTTTTCAACAAAGTTA	1664	
QY	659	TAAAGAAATACACATTCGAAGCTATTTTAGACATGGCCACATATCCCTCTATGGTAGCTA	718	
DB	16848	TTAAAAGGGAACCTTTCAAAGTGGTTAAGATATGATAATCTATCATCTGTCTGTAGCTA	1658	
QY	719	CTTCTGTAGTTGGTATAGGACTTTTTCGAAAGTGGTGGGTGGAAATAAACTGAGTACAGAAA	778	

Db 16588 CTTGTGCTTCTAGTGGGACTTTTTCGAAAGTGGAGAGTGGAAAGGTTTGAAGAAAGAAA 16529  
QY 779 TGAAGAGTTTCAACTAGGGAAGAGCTCATACATATTTTGTATATAAATCATCGGTTCAACGATAT 838  
Db 16528 TGGAGGGTATAGTTGGGTAAAGCTTCTATCTAGTACTGACTCTGACTTTTACAGCTATAC 16469  
QY 839 CATGGCAAGCTTGTGATTTGGAAGTGTGGTTTGAATATCGAAGTTTCAATCGCTTTT 998  
Db 16468 TTTGGCAAGTCCTTAGTATTTGGTAGTGGGACTGATTAATTTAGGCTCTCGCACTCTTCT 16409  
QY 899 CCAATGCTAAGCACTCTTTTGTATACAGTGTGCTGTCTTGTGTTGTCTTCTTCTCC 958  
Db 16408 CTAATGCTATAGCACTTTGGGGTGGCTTAATCTCTATATTGGCTGTGATGTTCTTTC 16349  
QY 959 GTGATGAGATGAGTGGAAATCAAGTTGGTGGCAATGTTTGGCCATCTGGGATTTGTTT 1018  
Db 16348 ATGACAGATGATGGAATCAAGATTATCTCTATGTTGGTGTGCTGCTGGGATTTATAT 16289  
QY 1019 CTTATGTTATCAGCATTTATGTCATGATAGAAAGCCAGAGAAAG 1063  
Db 16288 CTTATGTTATCAGCAGTACTCTGGATGAATGCAAGTCCCAACAG 16244

## RESULT 6

AF370622  
LOCUS AF370622.1 1106 bp mRNA linear PLN 30-APR-2001  
DEFINITION Arabidopsis thaliana putative protein (T9A21.60) mRNA, complete cds.

## ACCESSION

AF370622

## VERSION

AF370622.1 GI:13877726

## KEYWORDS

FLI\_CDNA.

## SOURCE

Arabidopsis thaliana (thale cress)

## ORGANISM

Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
1 (bases 1 to 1106)

## REFERENCE

Lam, B., Southwick, A., Karlin-Neumann, G., Nguyen, M., Miranda, M.,  
Palm, C.J., Bowser, L., Jones, T., Banth, J., Chen, H., Cheuk, R.,  
Chung, M.K., Kim, C., Lin, J., Liu, S.X., Pham, P.K., Sakano, H.,  
Shinn, P., Yamada, K., Ecker, J., Theologis, A. and Davis, R.W.  
Direct Submission

## JOURNAL

Submitted (17-APR-2001) DNA Sequencing and Technology Center,  
Stanford University, 855 California Avenue, Palo Alto, CA 94304,  
USA

## COMMENT

e-mail for correspondence: arab@sequence.stanford.edu

## FEATURES

This clone was isolated by RT-PCR.

## source

1. .1106  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
/chromosome="IV"  
/clone="C531"  
/ecotype="Columbia"  
1. .1106  
/gene="T9A21.60"  
657. .1106  
/gene="T9A21.60"  
/codon\_start=1  
/product="putative protein"  
/protein\_id="AAK43941.1"  
/db\_xref="GI:13877727"

## Gene

## CDS

translation="MDMIYVSLVASCVSUWGLFASSEWKLTSSEMDNYKHGKVSIM  
NLVATVWQVPSIGTGLIFELSLFSAISVGLPVPVPIIAVIFPHDKMGLKVIS  
MILAIWFTSYVYQQFLDNLKLNKHEITTESPPPEAEBSWQSK"

## ORIGIN

Query Match 28.0%; Score 320.6; DB 8; Length 1106;  
Best Local Similarity 57.9%; Pred. No. 5.2e-67;  
Matches 616; Conservative 0; Mismatches 429; Indels 18; Gaps 2;

QY 57 TCACAAAGTACTCTTGAGGAGTTAAGAGTGTCTCTCTATGTCACCTCTCTCTAGCTGGAGA 116

RESULT 7

Db 47 TAACACATACAAACGCTGGCTCAGGCTGACTTTATATACATCTCTTGTCAATTCAGGCCA 106  
QY 117 GACAATAGCCACTCTCTTAGGTAGACTTTTACTACGAAAGGCGGTAAAGCAGCATGGCT 176  
Db 107 AACAGTTGCTACAAATTTTGGGAGAGTATACTATGACACGGAGAAACAGTAATATGGCT 166  
QY 177 CGAAACCTTTGGTTCAGCTTTGTAGGGTTTCTTTTAAACCTTTCCTTGTCTATTTACTTAAA 236  
Db 167 AGCAAGGTTAGTTCAACTTTTGGCTTTCTCTGCTACTTCCATATATATCTTGTCAATT 226  
QY 237 GCTGAGCGCTCAAGACTAAACCAATTAACGAAAAAACTACTTCTTCTCTTGTGACACT 296  
Db 227 TAAAAACACATGCAACAACCTGATAGAGATGGAAGAAAGAACCTCACCT-----AGGAACCG 280  
QY 297 ATCTTTAGTGTATATTTGGACTTTGGCTTGTCTGCTGAGACATTTGATTTTGTACTCAATT 356  
Db 281 TGTATTTGGTTACGTAGTGTGGACTTCTTGTAGAGCAGATTTGCTATCTGTACTCCAT 340  
QY 357 TGGGCTACTTTACCTTCTCTCTCAACTTTTCTTTTGTATCTCTGGTTCGCAATTTGGCTTT 416  
Db 341 TGGACTTCTTTACTTACCCTTTTCTACCTATTTCCCTGATCTCTGTCATCTCAGTTAGCCTT 400  
QY 417 TAAACGCGCTTCTCTCTTACTTCTTAACTCAACAAAAATCAACANATTTATATCTCAATT 476  
Db 401 CAATGCTTCTCTCTCTTATTTTCTTAACTCAACAAAACTTACCCCTATCAITTTTAAATTC 460  
QY 477 ACTTGTCTCTTAAACCATATCTTCTACACTTCTGTTATCCAAACATGAACACAGATCTCC 536  
Db 461 TCTTTTCTCTTAACTATATCTTCCACCTTCTTCAATTTCAATTAAGAGACAGACTC 520  
QY 537 CTCTCTTACTTCAAAAGTCCGACGCAAGTCCAAAGTATGTGATTTGGATACATCTGGCGGT 596  
Db 521 CACAAAAGTTACAAAAGGAG-----AGTATGTCAAAGGTTTCATATGACCGT 568  
QY 597 CGGTAGCTCAGCTGGTATTTCTCTGGTGTCTTCTTTTAAACAGATTTAGCGGTTCCGAAAGAT 656  
Db 569 TCGTCTGCTCTGCTGGTTATGTTCTTATCTTATCCCTACAAACAGTAGCCCTTTCTTAAAGT 628  
QY 657 TCTAAGAAATACACATTTCAAGGCTTTTATAGACATGGCCACATATCCGCTCTATCGTAGC 716  
Db 629 CCTAAGAAAGCAAAATTTCTCAGAGTTATGATATGATAATCTACGTAGTCTAGTGGC 688  
QY 717 TACTTGTGTAGTTGTGGTAGGACTTTTGGAGTGGTGGGTGGAAAAAGCTGAGTACAGA 776  
Db 689 CAGTTGTGTAGCGTGGTGGGCTTTTGTCTAGCAGTGAAGTGAAGAACTTTGAGCAGTCA 748  
QY 777 AATGAAGAGTTTCAACTAGGAAAGCTCATACATTTTGTATGATAAACATCGGTTCAACGAT 836  
Db 749 AATGGATAACTACAAACATGGGAAGGTATCTTACATTTATGAACCTAGTGTGGACAGCTGT 808  
QY 837 ATCATGGCAAGCTTCTTTGATTGGAAGTGTGGTTTGTATTCGAAGTTTTCATCGCTTTT 896  
Db 809 TACCTGGCAGGATTTCTCATCGGTGGCAGGAGTGAATCTTCAGAGCTCTCTCTCTATT 868  
QY 897 TTCCAATGTCTAAGCACTCTTTTGTATACAGTGTGCTGTGTTGTCTTGTCTTCTTCT 956  
Db 869 CTCAAATGCAATGAAGGTTTGGGACTCCAGTGGTTCCTATCTTGGCTGTAAATCATTTT 928  
QY 957 CGGTATGAGATGAGTGGAAATCAAGTTGGTTGCAATGTTTTTGGCCATCTGGGGATTTGT 1016  
Db 929 CCATGACAAAATGAATGGGTTAAAGGTGATTTCTATGATCTCTAGCTATTTTGGGGTTTCA 988  
QY 1017 TTCTTATGTTATCAGCATATGTCATGATAGAAAGCCAGAGAGACCAAGAGCTTCC 1076  
Db 989 TTCTATGTTCTACCAACAAATATCTTGTATGACAAAAAATCTTGAAGAAAAATCATGAATCAC 1048  
QY 1077 TCAGTCTTAAAGAAAGAAAGAAACAAAAACAAGTAGATACCAATT 1119  
Db 1049 AACACAGATATCCCTTGACCCACAGAGAGAGAGTCAACT 1091



Db 898 GGATTAATCTTCGAGTCTTCTCTGTTCTTCCAAATTCCTCAATCTGCTGGGACTGCCT 957  
Qy 928 GTGTGCGCTGTTCTTCTGCTGTTCTTCTCCGTCGATCAGATGAGTGAATCAATGTTGTT 987  
Db 958 ATAGTTCCAGTTGAGCAGTGAAGTATTTTCCATGATGAAGATGATGATCAAAAGATTTTC 1017  
Qy 988 GCAATGTTTTTGGCCATCTGGGATTTGTTTCTTATGTTTATCAGCATTATGTCATGAT 1047  
Db 1018 TCCATCATTTTATAGTATCTGGGCTTCCCTTTTCATTTGCTATCAGCACTACCTCGACGA 1077  
Qy 1048 AGAAG 1053  
Db 1078 AAGAAG 1083

RESULT 8  
AY074546  
LOCUS  
DEFINITION  
Arabidopsis thaliana unknown protein (At4g18200) mRNA, linear PLN 18-SEP-2002  
cds.

ACCESSION  
AY074546  
VERSION  
AY074546.1 GI:18491220  
KEYWORDS  
FLU CDNA.  
SOURCE  
Arabidopsis thaliana (thale cress)  
ORGANISM  
Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eudicots II; Brassicales; Brassicaceae; Arabidopsis.  
REFERENCE  
Yamada, K., Liu, S.X., Sakano, H., Pham, P.K., Banh, J., Chung, M.K.,  
Goldsmith, A.D., Lee, J.M., Quach, H.L., Toriumi, M., Yu, G., Bowser, L.,  
Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J.,  
Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B.,  
Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T.,  
Satou, M., Seki, M., Shinn, P., Southwick, A., Shinozaki, K.,  
Davis, R.W., Ecker, J.R. and Theologis, A.  
Arabidopsis Full Length cDNA Clones

Unpublished  
TITLE  
Arabidopsis Full Length cDNA Clones  
JOURNAL  
2 (bases 1 to 1428)  
REFERENCE  
Yamada, K., Banh, J., Chan, M.M., Chang, C.H., Chang, E., Dale, J.M.,  
Deng, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L.,  
Tang, C.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Bowser, L.,  
Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J.,  
Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B.,  
Lin, J., Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J.,  
Sakurai, T., Sacou, M., Seki, M., Shinn, P., Southwick, A.,  
Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.  
Direct Submission  
Submitted (16-JAN-2002) Plant Gene Expression Center, 800 Buchanan  
Street, Albany, CA 94710, USA  
COMMENT  
RIKEN Genomic Sciences Center (GSC) members carried out the  
collection and clustering of RAPL cDNAs (RAPL cDNA : 'RIKEN  
Arabidopsis Full-length cDNA') : Seki, M., Narusaka, M., Ishida, J.,  
Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,  
Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PGENC (SSP) Consortium members carried out the  
sequencing and annotation of the RAPL cDNAs: Yamada, K., Banh, J.,  
Chan, M.M., Chang, C.H., Chang, E., Dale, J.M., Deng, J.M.,  
Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C.,  
Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Bowser, L., Chen, H.,  
Cheuk, R., Jones, T., Karlin-Neumann, G., Kim, C., Lam, B., Lin, J.,  
Meyers, M.C., Miranda, M., Nguyen, M., Palm, C.J., Shinn, P.,  
Southwick, A., Davis, R.W., Ecker, J.R. and Theologis, A.

Yamada, K. (SSP/PGENC) and Seki, M. (RIKEN GSC) contributed equally to  
this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP/PGENC)  
contributed equally to this work as PIs.

Annotation is based on the January 2002 version of the Arabidopsis  
genome submitted to GenBank.  
Location/Qualifiers

source  
1. .1428  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
/chromosome="4"  
/clone="RAFL09-36-N15 (R18628)"  
/ecotype="Columbia"  
/note="This clone is in a modified pBluescript vector  
(FLC-1) as a BamHI/XhoI insert."  
gene  
1. .1428  
/gene="At4g18200"  
1. .71  
/gene="At4g18200"  
72. .1205  
/gene="At4g18200"  
/codon\_start=1  
/evidence=experimental  
/product="unknown protein"  
/protein\_id="AA169512.1"  
/db\_xref="GI:18491221"  
translaton="MGISQVHCNGDQNLLEALLDHEETESFSPOTKNCRWLRVSI  
YAFVFCQPLATVGLRYLGYENGKSTYVTVLLQLIGPVLILFRFSRIPKSDYT  
NFSQSPFTLASVYLCTLLSAYALSAGVLLLPVSTLSLILFRFSRIPKSDYT  
LNSQKPTPLIVNSFLLTSSALLVNDSENTNVRVYVIGFTCTTGASAGILV  
LSLIQLPRKFTKTSGLVLDANYQSLVLCVVLIGLFGSEWRTLFSENRKYLK  
KYSYILTASAAIFWQVTVGCVGLIFESSVSFNSITAVGLPIVPPVAVIVPHDKMD  
ASKIFSIILAIWGLSFYVQYHLDKLUKTCQTKPVEBETQL"  
1206. .1428  
/gene="At4g18200"  
3'UTR  
misc\_difference 1412  
/gene="At4g18200"  
/note="not present in genomic sequence"  
ORIGIN  
Query Match 27.8%; Score 318.4; DB 8; Length 1428;  
Best Local Similarity 57.8%; Pred. No. 1.7e-66;  
Matches 593; Conservative 0; Mismatches 421; Indels 12; Gaps 1;  
Qy 28 GAAGGGAAATTTTCAACAGAGAGAGAGTCAAGTCTCTTGGAGGTTAAGAGTGCT 87  
Db 141 GAACTGAATCAATTTTCAGTACTCTCAACAGAGAACTGTAAGAGGTGGCTCGGTGCTCC 200  
Qy 88 CTCATGTGCTCTCTCTTCTAGCTGGAGAGACAATAGCCACTCTCTTAGGTAGACTTTAC 147  
Db 201 ATATAGCAATCTTCTGTCATCTTCTGCCAACCACTGCTACAGTTCTGGGTAGACTGTAC 260  
Qy 148 TAGCAAAAAGGGGTAAGAGACATGGCTCGAAACCTTGGTTCAGCTTGTAGGTTTCCT 207  
Db 261 TATGAAAATGGAGGAAAAGACATATGTGGTAACACTTCTTCAACTCATTTGGCTTCCCT 320  
Qy 208 TTAACCCCTTCTGCTATTATTACTTAAGCCGTGAGCGCTCCAAAGACTAAACATTACC 267  
Db 321 GTACTGATTTCTTCCGCTCTTCTTTCGAAATCAGCAACCAAAATCAACAGATCAAAAT 380  
Qy 268 AAAAAAACTACTTCTTCTTCTTGACACTATCTTTAGTGTATATTGGACTTGGCTTGCTT 327  
Db 381 TTCAGTCAGTCCCTTCTTCCACCCTTGCATCGTTTACTTGTGCACTGACTGCTA 440  
Qy 328 GTTGTGGACATGTATTGTTGACTCATTTGGGCTACTTTGCTGCTCAACTTTC 387  
Db 441 GTGTCTGCTTATGCTTATTGTTCTGCAGTTGGGTTGCTTTACTTACCACTCTCTACTTTC 500  
Qy 388 TCTTTGATCTCTGCGTGGCAATTTGGCTTTTAACGGCTCTTCTTACTTCTTAACTCA 447  
Db 501 TCCCTCATCTTGGCTTCAAGTTGGCTTCTCATGCTTTTCTCATATTTCTTAACTCG 560  
Qy 448 CAAAAAATCACCACTTTTATCTCAATTCACCTTGTCTTCTTAACCATATCTTCTACACTT 507  
Db 561 CAAAGTTCACTCTTGTGATAGTCAATCTTGTCTTCTTCTTACGGTTTCTCTGCTCCTC 620  
Qy 508 CTTGTTATCCAAATGAAACCAAGATCTCCCTCTTCTTACTTTCAAAGTCCGAGCAAGTCC 567  
Db 621 CTCGTGGTCAACACTGATTCAGAA-----AACACAACAAATGTATCTAGAGTA 668

```
QY 568 AAGTATGTGATTGGATACATCTGCGGGTCGGTAGCTCAGCTGGTTATTCTCTGGTGCTT 627
Db 669 CAATATGTGATTGGGTTCACTGTACCATTTGGTGTCTCCGTGGGATGGATTTGGTACTA 728
QY 628 TCTTTTAACAGATFACCGGTTCCGAAAGATTTCTAAAGAAATACACATTCAAGGCTATTTTAA 687
Db 729 TCTCTGATACAACTGCTCTTCAGGAAAGTTTTCAGGAAGCATACATCTCAGCAGTCCTG 788
QY 688 GACATGGCCACATATCCGTCTATGTAGTACTTGTGTAGTTGTGGTAGGACTTTTGGGA 747
Db 789 GACITGGCCAAATFACCAAGTCTCTAGTTGCAACTTGTGTGACTCATAGGACTGTTTGCA 848
QY 748 AGTGTGGGTGGAAGAGCTGAGTACAGAAATGGAAGAGTTTCAACTAGGGAAGCTCA 807
Db 849 AGTGAGAGTGGAGAACTCTGCCAAGTGAAGATGAGAACTCAAACTGGGGAAGTGTC 908
QY 808 TACATTTTGTATAAATCATCGGTTCAACGATATCATCGCAAGCTTGTGTTGATGGAAGTGT 867
Db 909 TATATCTTTGACTTTGGCCCTCAGCAGCTATTTCTGGCAAGTATACACATGTTGGTTGGTG 968
QY 868 GGTTTGATATCGAAGTTTCACTCGCTTTTTCGAATGTCATAAGACATCTTTGTTTACCA 927
Db 969 GGAATTAATCTTCGAGTCTTCTCTGTGTCTTCCAAATCCAAATCCAAATCCAAATCCAA 1028
QY 928 GTTGTGCTGTTTCTTGTGTTGTTCTTCTCGTGATGAGATGAGTGAATCAAGTTGGTT 987
Db 1029 ATAGTTCAGGTTGAGCAGTATAGTTTTCATGATGAAGATGAGTGAATCAAGATTTTC 1088
QY 988 GCAATGTTTTTGGCCATCGGGATTTGTTTCTTATGTTTATCAGCAATATGTCAATGAT 1047
Db 1089 TCCATCATTTTAGCTATCTGGGCTTCTCTCAITTTGTTCTATCAGCACTACCTCGACGA 1148
QY 1048 AGAAG 1053
Db 1149 AAGAAG 1154
```

```
RESULT 9
AX507517 AX507517 3387 bp DNA linear PAT 27-SEP-2002
LOCUS Sequence 2212 from Patent WO0216655.
DEFINITION AX507517
ACCESSION AX507517
VERSION AX507517.1 GI:23388754
KEYWORDS Arabidopsis thaliana (chale cress)
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
REFERENCE 1
AUTHORS Harper, J.F., Krops, J., Wang, X. and Zhu, T.
TITLE Stress-regulated genes of plants, transgenic plants containing
same, and methods of use
JOURNAL Patent: WO 0216655-A 2212 28-FEB-2002;
The Scripps Research Institute (US) ; Syngenta Participations AG
(CH)
FEATURES
source Location/Qualifiers
1..3387
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"
ORIGIN
Query Match 27.8%; Score 318.4; DB 6; Length 3387;
Best Local Similarity 57.8%; Pred. No. 1.6e-66;
Matches 593; Conservative 0; Mismatches 421; Indels 12; Gaps 1;
QY 28 GAAGGGAATTTTCAACAGAGAGAGAGTACCAAGTACTCTTGGAGGTTAAGAGTCTCT 87
Db 2323 GAAACTGAATCAITTTTTCAGTACCTCAACGAGAACTGTAAAGAGTGGCTCCGTGTCTCC 2382
QY 88 CTCTATGTCACTCTCTCTTTAGCTGGAGAGACAATAGCCACTCTCTTTAGGTAGACTTTAC 147
```

```
Db 2383 AFATACGAATCTTTGTGTCATCTTCTGCAACCACTTGTACAGTTCTGGGTAGACTGTAC 2442
QY 148 TACGAAAAGCGGTAAAGACATGGCTCGAAACCTTTGGTTTCAAGCTTTGTAGGTTTCCCT 207
Db 2443 TATGAAAATGAGGGAAGACATATGTGTAACACTTCTTCAACTCATTTGGCTTCCCT 2502
QY 208 TTAACCCCTTCTCTCTATTTACTTAAAGCCTGAGCGGTCCAAAGACTAAACCAATACC 267
Db 2503 GTACTGATTTCTTCTCCGCTTCTTTCTCGAATCAGGCAACCCAAATCAACAGATACAAAT 2562
QY 268 AAAAAAATCACTCTTCTTCTTCTGACACTATCTTTTAGTGATATATTGGACTTTGGCTTGCTT 327
Db 2563 TTCACTCAGTCCCCTTCTTCTCACCACTGTCATCGGTTTACTTGTGCATGGAGTGTCTA 2622
QY 328 GTTCTGAGCAATGTATTTTGTACTCATTTGGGCTACTTTACCTTCTCTGTCTCAACTTTC 387
Db 2623 GTGTCTGTTATGCTTATTTGTCTGCACTGGGTTGGTTTACTTACCAGTCTCTACTTTTC 2682
QY 388 TCTTTGATCTCTGGCTCGCAATTTGGCTTTTAAAGCGCTTCTCTTACTTCTTAAACTCA 447
Db 2683 TCCCTCATCTTGGCTCACAGTTGGCTTCACTGCCCTTTTCTCATATTTCTTAACTCG 2742
QY 448 CAAAAAATCACCAATTTTACTCAATTCATCTTCTTCTTAAACCATATCTTCTACACTT 507
Db 2743 CAAAAGTTCACTCTTGTGATGATCAATCTTTTGTCTTCTCTTACGGTTTCTCTGCCCTC 2802
QY 508 CTTGTTATCCAAATGAACAGCAATCTCCCTTCTTACTTCTTCTTCTTCTTCTTCTTCTT 567
Db 2803 CTCGTGGTCAACACTGATTGAGAA-----AACACAACAATGATCTAGAGTA 2850
QY 568 AAGTATGTGATTTGATATACATCTGCGCGGTGCGTAGCTCAGCTGGTTTATTTCTCTGGTCTT 627
Db 2851 CAATATGTGATTTGGGTTTCACTCTGTACCAATTTGGTGTCTTCCGTGGGATTTGGATTG 2910
QY 628 TCTTTAACAGATTACGGGTTTCGAAAAGATTCTAAGAAATACACATTTCAAGGCTATTTTA 687
Db 2911 TCTCTGATACAACTGCTCTTCAGGAAGTTTTCAGGAAGCATATCTCTCAGCAGTCTG 2970
QY 688 GACATGCCACATATCCGTCTATGGTAGTACTTGTGTAGTGTGTGGTAGGACTTTTGGGA 747
Db 2971 GACTTGGCCAAATACCAGTCTCTAGTTGCAACTTGTGTGTACTCTATAGGACTCTTTGCA 3030
QY 748 AGTGTGGGTGGAAGAGCTGAGTACAGAAATGGAAGTTTCAACTAGGGAAGCTCA 807
Db 3031 AGTGAGAGTGGAGAACTCTGCCAAGTGAGATGAGAAACTACAAACTGGGGAAGTGTCA 3090
QY 808 TACATTTTGTATAAATCATCGGTTCAACGATATCATGGAAGCTTGTGTTGGAAGTGT 867
Db 3091 TATATCTTGACTTTGGCTCAGCAGCTATTTTCTGGCAAGTATACACTGTTGGTTGTG 3150
QY 868 GGTTTGATATCGAAGTTTCAATCGCTTTTTCGAATGTCTATAAGCACTCTTTGTTTACCA 927
Db 3151 GGATTAATCTTCGAGTCTTCTTCTGTGTTCTCCAAATTCATAACTGCTGTGGGACTGCT 3210
QY 928 GTTGTGCTCTTCTTGTCTGTGTTCTTCTCGGTAGTGAAGTGAATCAAGTGTGTT 987
Db 3211 ATAGTTCAGTTGTAGCAGTGTAGTTTTCATGATAAGATGATGATCAAGATTTTC 3270
QY 988 GCAATGTTTTTGGCCATCTCGGATTTGTTTCTTATGTTTATGAGCAATATGTCAATGAT 1047
Db 3271 TCCATCATTTTAGCTATCTGGGCTTCTTTCATTTGTTCTATCAGCACTACCTCGACGA 3330
QY 1048 AGAAG 1053
Db 3331 AAGAAG 3336
RESULT 10
AX652015
LOCUS Sequence 908 from Patent WO0300098.
DEFINITION AX652015
ACCESSION AX652015
```



```

VERSION      AX652015.1  GI:29154833
KEYWORDS
SOURCE
ORGANISM      Arabidopsis thaliana (thale cress)
               Arabidopsis thaliana
               Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
               Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
               rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE
AUTHORS      Chang, H.S., Chen, W., Cooper, B., Glazebrook, J., Goff, S.A., Hou, Y.M.,
               Katagiri, F., Qian, S., Tao, Y., Whitham, S., Xie, Z., Zhu, T. and Zou, G.
TITLE        Plant genes involved in defense against pathogens
JOURNAL      Patent: WO 0300898-A 908 03-JAN-2003;
               Syngenta Participations AG (CH)
FEATURES
source       1..3387
               /organism="Arabidopsis thaliana"
               /mol_type="unassigned DNA"
               /db_xref="taxon:3702"
ORIGIN
Query Match      27.8%; Score 318.4; DB 6; Length 3387;
Best Local Similarity 57.8%; Pred. No. 1.6e-66;
Matches 593; Conservative 0; Mismatches 421; Indels 12; Gaps 1;

QY 28 GAAGGGAATTTTCAACAGAGAGAGAGTCAAGTCACTCTTTGGAGGTTAAGAGTCT 87
DB 2323 GAAACTGAATCAATTTTCAGTACCTCAACAGAGAGAGTCAAGAGTCTCTCC 2382
QY 88 CTCATGTCACTCTCTCTAGCTGAGAGACATAGCCACTCTCTTAGTACATTTAC 147
DB 2383 ATATACGCAATCTTTGTCACTCTCTCCCAACCACTTGTACAGTCTCTGGGTAGATGATAC 2442
QY 148 TAGCAAAAAGGGGTAAAAAGCACATGGCTCGAAACCTTGGTTCAGTCTTAGGGTTTCC 207
DB 2443 TATGAAATGGAGGGAAGACATATGTGGTAACACTTCTTCAACTCACTAGTCTCTCT 2502
QY 208 TTAACCCCTCTCTGCTATTATTAATTAAAGCCGTGAGCCGTGCAAGACTAAACCAATTACC 267
DB 2503 GTACTGATTTCTCTCCGCTCTTTCTCGAATCAGGCAACCCCAATCAACAGATACAAT 2562
QY 268 AAAAAAATCTCTCTCTCTGACACTATCTTTAGTGTATATTTGGACTTGGCTTGTCT 327
DB 2563 TTCAGTCACTGCTCTCTCTACCACTTGTGATCGTGTACTTGTGCACTGAGCTGCTA 2622
QY 328 GTTGTGCAATGTATTTGTACTCATTTGGGCTACTTTAGCTTCTCTGCTCAACTTTC 387
DB 2623 GTGCTGCTATGCTATTATTTGCTGAGTTGGTGTCTTACTTACCAGTCTCTACTTTC 2682
QY 388 TCTTTGATCTCTGCGTGGCAATTTGGCTTTTAAGCGCGTCTCTCTTACTTCTTAAACTCA 447
DB 2683 TCCCTCATCTTGGCCTTCAAGTTGGCTTCACTGCGCTTTTCTCATATTTCTTAACTCG 2742
QY 448 CAAAAATCACACCAATTTATCTCAATCTGTTCTTCTTAAACCATATCTTCACTT 507
DB 2743 CAAAGTTCACCTCTTGTAGTCAATCTTTGTTTCTCTTACGGTTTCTCTGCGCTC 2802
QY 508 CTTGTTATCCAGATGAACCAAGATCTCCCTCTCTTACTTTCAAGTCCGACCAAGTCC 567
DB 2803 CTCGTGCTCAACACTGATTCAGAA-----AACACAAACAATGATCTAGAGTA 2850
QY 568 AAGATGTGATTTGATATATCTGCGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 627
DB 2851 CAATATGTGATTTGGTGTCTGATACCATTTGGTGTCTTCCGCTGGGATTTGGATTTGACTA 2910
QY 628 TCTTTAACAGATTACCGGTTCGAAGATTTCTTAAGAAATACACATTCAAGCTATTITA 687
DB 2911 TCTCTGATACAACTGCTCTTCAAGGAAAGTTTTCAGGAAGCATACATCTCAGCAGTCTCG 2970
QY 688 GACATGGCCACATATCCGCTCTAGTGTAGTCTTGTGTAGTGTGTGTAGTGTGTGTGTGA 747
DB 2971 GACTTGGCCAAATACAGTCTCTAGTGTGAATTTGTGTGTGTGTGTGTGTGTGTGTGA 3030
QY 748 AGTGGTGGTGGAAAAAGCTGAGTACAGAAATGGAAGAGTGTTCCTTCACTAGGGAAGAGCTCA 807
```

```

DB 3031 ACTGGAGAGTGGAGAACTCTGCCAAGTGAAGATGAGAAACTGCGGAAAGTGTCA 3090
QY 808 TACATTTTGTATAAATCATCGTTTCAACGATATCATGCAAGCTTGTGTGATTCGAACTGTT 867
DB 3091 TATATCTTGAATTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGG 3150
QY 868 GGTGTTGATTTATCAAGAGTCTTTCATCGCTTTTTCCTCAATGTATCAAGCACTCTTTGTTTACCA 927
DB 3151 GGATTTAAATCTTCGAGTCTTCTTCTGTTCTTCAATTCATTAATTCATGCTGCTGGACTGCT 3210
QY 928 GTTGTGCTGCTGTTCTTCTGCTGTTCTTCTTCTGATGATGATGATGATGATGATGATGATGAT 987
DB 3211 ATAGTTCCAGTTGTAGCAGTGTATGTTTTCATGATGATGATGATGATGATGATGATGATGAT 3270
QY 988 GCAATGTTTTCGCGCATCTGGGATTTGTTCTTATGTTATGATGATGATGATGATGATGATGAT 1047
DB 3271 TCCATCATTTTAGCTATCTGGGCTTCTTTCATTTGTTCTATCAGCACTACCTCGACGAA 3330
QY 1048 AGAAAG 1053
DB 3331 AAGAAG 3336

RESULT 11
AX033548
LOCUS
DEFINITION   Sequence 5 from Patent WO049152.
ACCESSION   AX033548
VERSION     AX033548.1  GI:10280292
KEYWORDS
SOURCE       Arabidopsis thaliana (thale cress)
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE
AUTHORS      Andre, B., Buerkle, L., Frommer, W.B. and Gillissen, B.
TITLE        Nucleic acids that code for a nucleobase transporter
JOURNAL      Patent: WO 0049152-A 5 24-AUG-2000;
               ANDRE BRUNO (BE); BUERKLE LUKAS (DE); FROMMER WOLFF B (DE);
               GILLISSEN BERND (DE)
FEATURES
source       1..1194
               /organism="Arabidopsis thaliana"
               /mol_type="unassigned DNA"
               /db_xref="taxon:3702"
ORIGIN
Query Match      27.7%; Score 317.2; DB 6; Length 1194;
Best Local Similarity 57.5%; Pred. No. 3.4e-66;
Matches 596; Conservative 0; Mismatches 428; Indels 12; Gaps 1;

QY 18 GGTAGAACCTGAAGGAAATTTTCAACAGAGAGAGAGTCAAGTACTCTTTGAGGTT 77
DB 75 GGTGTAATCTGAATCATCATCATCATGCTGTGCTCAAAACCGAGAACTATAAAAGTGGCT 134
QY 78 AAGAGTGTCTCTATGTCACTCTCTCTAGCTGAGAGACAATAGCCACTCTCTTAGG 137
DB 135 TCGTGTCTCCATATACGTAATCTTTGTCTCTTTTGGCAGGCACCTAGCTACAAATCTCTGG 194
QY 138 TAGACTTTTACTACGAAAAAGCGGTAAAGCACATGCTCGAAACCTTTGGTTTCACTTGT 197
DB 195 TAGATTGACTATGAATGGAGGAATAGCACATATGTGTTAACTCTTCTCACTCAT 254
QY 198 AGGGTTTCTCTTAACCCCTTCTCTGCTATTTATTTAAAGCGTGAAGCTGCAAGACTAA 257
DB 255 TGGCTTCCCTGCTACTGTTCTGTTCCGCTTCTTTTCTCGAATCAGGCAACCCAAATCAAC 314
QY 258 AACCATTTACAAAAAATCACTCTTCTCTCTTCTTGACATATCTTTAGTGTATATTGGACT 317
DB 315 AGATACAAATTTTCACTAGTCCCTTCTTCCACCACTTCTGATCGGTTTACTTGTGCAC 374
```

Qy	318	TGGCTTGCTGTTGTCGGACATGTTATTTTGTACTCAATTTGGGTACTTTTACCTTCCTGT	377
Db	375	TGGACTGCTAGTGTCCGCTTATGCTTATTTGTGTCAGTAGGGTGTCTACTTACCAGT	434
Qy	378	CTCAACTTTCCTTTGATCTCTGGTGCACAAATGGCTTTTAAACGCCGCTTCTCTTACTT	437
Db	435	CTCTACTTTCCTCCCTCATCTTGGCTCACAGTTGGCTTCACTGCCCTTTTCTCATATTT	494
Qy	438	CCTAAATCACAAAAAATACACCAATTATATCTCAATTCATCTTGTTCCTTAAACATATC	497
Db	495	CCTTAACTCGCAAAAGTTCACTCCTTTGATAGTCAGTTCCTTGTCTCTCCTCACTGATC	554
Qy	498	TTCTACACTTCTGTTATCCAAATGAACACAGATCTCCCTCTTCTACTTCAAAAGTCGC	557
Db	555	CTCTGCTCTTCTGTGGTCAACACTGATTCAGAA-----AACTCAACTAATGT	602
Qy	558	AGCCAAGTCCAAGTATGTATGGATACATCTGGCGGTGCGGTAGCTCAGCTGGTGTATTC	617
Db	603	ATCTAGAGTACAGTATGTATCGGGTTCATATGTACCATCGGTCTCCGCTGGGATGG	662
Qy	618	TCTGGTGTCTTTTAAACAGATTACGGCTTCGAAAGATTCCTAAAGAAATACACATCAA	677
Db	663	ACTGTTACTATCTCTGATACAAATGCTCTTCAGGAAAGTTTTCAGGAAGCATACATCTC	722
Qy	678	GGCTATTTTACACATGGCCACATATCCGCTCTATGGTAGTACTTGTGTAGTTGTGGTAGG	737
Db	723	AGCAGTACGGACTTGGCCATTTACCAGTCTCTAGTTGCGAGTTGTGTAGTTCTCATAGG	782
Qy	738	ACTTTTTCGAAAGTGTGGGTGGAAAAAGCTGAGTACAGAAATGGAAGAGTTTCAACTAGG	797
Db	783	ACTTTTTCGAAAGTGGAGAGTGGAAACCTTGCCAAAGTGAGATGAGAACTACAACTCGG	842
Qy	798	GAAGAAGCTACATATTTTGAATTAACATCGGTTTCAACGATATCATGGCAAGCTTGTGTAT	857
Db	843	GAAGAAGTGCATATGTTTGAATTTAGCTTGGCCGCGAGTATTTCCGCGCAAGTCTACACT	902
Qy	858	TGGAAGTGTGCTTGTATTCGAAAGTTTTCATCGCTTTTCCAAATGTCATAAGCACTCT	917
Db	903	TGGTCTTGGGAATGATCTTCGAGTCATCTCTGTGTCTTCCAAATCCATAACAGCTGT	962
Qy	918	TGTTTACAGTGTGCTGTTCTTGTCTGTGTCTTCTTCCGTCATGAGATGAGTGAAT	977
Db	963	GGGATGCTATAGTTCAGTTGCGGAGTATAGTATTTCCATGATAGATGGAGCATC	1022
Qy	978	CAAGTGGTGGCAATGTTTGGCCATCTGGCAATCTGGGATTTGTTCTTANGTTATACGATTA	1037
Db	1023	AAAAATCTTCCATATTTTATAGCTATCTGGGCTTCCCTTCAATTCGTCATCAGCACTA	1082
Qy	1038	TGTCAAATGATAGAAG	1053
Db	1083	CCTCGACGAAAGAAG	1098
RESULT 12			
AX033550			
LOCUS	AX033550	1071 bp	DNA linear PAT 21-SEP-2000
DEFINITION	Sequence 7 from Patent WO0049152.		
ACCESSION	AX033550		
VERSION	AX033550.1	GI:10280294	
KEYWORDS	Arabidopsis thaliana (thale cress)		
SOURCE	Arabidopsis thaliana		
ORGANISM	Arabidopsis thaliana		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;		
	rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi		
REFERENCE	1		
AUTHORS	Andre B., Buerkle L., Frommer W.B. and Gillissen B.		
TITLE	Nucleic acids that code for a nucleobase transporter		
JOURNAL	Patent: WO 0049152-A 7 24-AUG-2000;		
	ANDRE BRUNO (BE) ; BUERKLE LUKAS (DE) ; FROMMER WOLF B (DE) ;		
	GILLISSEN BERND (DE)		
FEATURES	Location/Qualifiers		
source	1..1071		

/organism="Arabidopsis thaliana"			
/mol_type="unassigned DNA"			
/db_xref="taxon:3702"			
ORIGIN			
Query Match 27.0%; Score 308.6; DB 6; Length 1071;			
Best Local Similarity 58.3%; Pred. No. 4.2e-64;			
Matches 589; Conservative 0; Mismatches 404; Indels 18; Gaps 2;			
Qy	111	TGGAGAGACAATAGCCACTCTCTTAGGTAGACTTTTACTACGAAAAAGCGGTAAAGCAC	170
Db	51	TGGCCAATCAGTTGCTACAATTCCTGGCAGACTATCTATGAAAAATGAGAGAAACAGCAA	110
Qy	171	ATGGCTCGAAACCTTGGTTGAGCTTGTAGGGTTCTTTAAACCTTCTCTGCTATTATTA	230
Db	111	ATGGCTAGCAACGGTAGTTCAGCTTGTAGGCTTTCTTATTTCTACTTCATATCATCTCT	170
Qy	231	CTTAAAGCCTGAGCGGTCCAAAGACTAAACCATTTACCAAAAAAACTACTTCTTCTCTT	290
Db	171	GTCAAGTCAAAACACATACAACAACTCAGAGAGATGGCAAAATTAA-----CCTCACTAG	224
Qy	291	GACACTATCTTTAGTGTATATTGGACTTGGCTTGTCTTGTCTGGACATGTAATTTGTA	350
Db	225	GAACCGTGCAATAGTTTACATAGTGTGGACTTCTTGTAGGAGCAGCTTGTACCTATA	284
Qy	351	CTCAATTTGGGCTACTTTACCTTCTCTGCTCAACTTCTTCTTCTTCTCTCTCGGTCGAAT	410
Db	285	TTCCATTTGSACTGCTTTTACCTGCTTCTTCTACCCCTTCCCTGATCTGTGCATCACAGTT	344
Qy	411	GGCTTTTAAACCGCTCTTCTTCTTACTTCTTAAACTCACAATAATCACACATTTACT	470
Db	345	AGCTTTCACCGCTTCTTCTTCTTACTCAACTCACAATAACTTACTCTCTATCATTTT	404
Qy	471	CAATTCATCTGTTCTCTTAAACCATATCTTCTCACTTCTTGTGTTATCCAAATGAACAGA	530
Db	405	GAATTTCTCTTCTCTCTCTCTACTATCTTCCACCTCTTGCATTTAATAACGAGGAA--	462
Qy	531	ATCTCCCTCTTCTTCTTCAAGTCCGACGACAGTCCAAAGTATGTAATGTAATGTAATCTG	590
Db	463	-----TCAGATTCCAAAAAAGTTACAAAGGAGAGTATGTCAAAGTTTCGTATG	512
Qy	591	CGCGTTCGCTAGTCTCAGCTGGTTATCTCTGGTGTCTTCTTTAAACAGATTACGCGTTT	650
Db	513	CACGTTGGTGCACTCTGCTGGGTTTGTCTACTCTTATCCCTACAACAGCTAGCTTTG	572
Qy	651	AAAGATTCTAAAGAAATACACATTCAAAGGCTATTTTAGACATGGCCACATATCCGCTAT	710
Db	573	TAAAGTTTAAAGAAAGCAAACTTTCTCAGAAGTTTATAAATATGATAATCTACATGAGT	632
Qy	711	GGTAGCTACTTGTGTAGTTCTGTTAGACATTTTGGAAAGTGGTGGTGGGAAAGCTGAG	770
Db	633	AGTGGCCAGTTGTGTAGCGTGGGGCTTTTGTAGTAGCGAGTGGGAAACTTTGAG	692
Qy	771	TACAGAAATGGAAGAGTTTCAACTAGGAAAAAGCTCATACATTTTGTATAAACATCGGTT	830
Db	693	CAGTGAATGGAAAACTACAAACTTGGGAAGGTATCCTATCTCATGAACCTAGTGTGAC	752
Qy	831	AACGATATCATGGCAAGCTTGTGTTGATGGAAAGTGTGGTTGATTTATCGAAGTTTCAT	890
Db	753	AGCTGTTACCTGGCAGGTAATTCCTCATCGGTTGACAGGACTGATCTTCGAGCTTCTCT	812
Qy	891	GCTTTTTCCAAATGCTATAGCACTCTTGTTTTACCAGTTGTGCTGTCTTCTGCTGTGT	950
Db	813	CCTATTTCTAAATGCAATAGCGCTTTTGGGACTCCCCGTGGTTCCTATCTCTGGCTGAT	872
Qy	951	CTTCTTCGCTGATGAGATGAGTGGAAATCAAGTTGGTTGCAATGTTTGTGGCAATCTGGG	1010
Db	873	CATTTTCCATGACAAAAATGAACGCTTAAAGGTGATTTCTATGATTCTAGCTATTTGGG	932
Qy	1011	ATTTGTTTCTTATGGTTATCAGCAATATGTCAATGATAGAAAGCCAGCAAGAACCAAGA	1070
Db	933	TTTCGTATCTATGCTTACCAACAATATCTTGTAGTAAACAACTTTGAAGAAAAAGTAATGA	992



Qy	1030	CAGCATTATGCAATGATAGAAAGCCAGAAAGACCAAGAGCTTCCTCA	1079	Query Match	26.6%	Score 304.8;	DB 8;	Length 1372;
Db	1063	CAGAATCATATTGATGACTTGAAGTAGAAGCAAGCAGCAACCAAGCTCA	1112	Best Local Similarity	58.1%	Pred. No. 3.5e-63;		
				Matches	587;	Conservative	0;	Mismatches 402;
							Indels	21;
							Gaps	2;
RESULT 14								
AY062782								
LOCUS	AY062782	1372 bp	mRNA	linear	PLN 25-NOV-2001			
DEFINITION	Arabidopsis thaliana Unknown protein (Atlg44750; T12C22.2)				mrna,			
complete cds.								
ACCESSION	AY062782							
VERSION	AY062782.1	GI:17065411						
KEYWORDS	FLI CDNA							
SOURCE	Arabidopsis thaliana (thale cress)							
ORGANISM	Arabidopsis thaliana							
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;							
AUTHORS	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;							
	rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.							
	1 (bases 1 to 1372)							
	Southwick,A., Karlin-Neumann,G., Nguyen,M., Lam,B., Miranda,M.,							
	Palm,C.J., Bowser,L., Jones,T., Banh,J., Carninci,P., Chen,H.,							
	Cheuk,R., Chung,M.K., Hayashizaki,Y., Ishida,J., Kamiya,A.,							
	Kawai,J., Kim,C., Lin,J., Liu,S.X., Narusaka,M., Pham,P.K.,							
	Sakano,H., Sakurai,T., Satou,M., Seki,M., Shinn,P., Yamada,K.,							
	Shinozaki,K., Ecker,J., Theologis,A. and Davis,R.W.							
	Direct Submission							
TITLE	Submitted (14-NOV-2001) DNA Sequencing and Technology Center,							
JOURNAL	Stanford University, 855 California Avenue, Palo Alto, CA 94304,							
	USA							
COMMENT	e-mail for correspondence: arabsequence.stanford.edu							
	RIKEN Genomic Sciences Center (GSC) members carried out the							
	collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN							
	Arabidopsis Full-length cDNA'): Seki,M., Narusaka,M., Ishida,J.,							
	Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,							
	Hayashizaki,Y. and Shinozaki,K.							
	The Salk, Stanford, PGEC (SSP) Consortium members carried out the							
	sequencing and annotation of the RAFL cDNAs: Nguyen,M.,							
	Southwick,A., Karlin-Neumann,G., Lam,B., Miranda,M., Palm,C.J.,							
	Bowser,L., Jones,T., Banh,J., Chen,H., Cheuk,R., Chung,M.K.,							
	Kim,C., Lin,J., Liu,S.X., Pham,P.K., Sakano,H., Shinn,P.,							
	Yamada,K., Ecker,J., Theologis,A. and Davis,R.W.							
	Southwick,A., (SSP/Stanford) and Seki,M. (RIKEN GSC) contributed							
	equally to this work. Shinozaki,K. (RIKEN GSC) and Davis,R.W.							
	(SSP/Stanford) contributed equally to this work as PIs.							
FEATURES	Location/Qualifiers							
source	1..1372							
	/organism="Arabidopsis thaliana"							
	/mol_type="mRNA"							
	/db_xref="taxon:3702"							
	/chromosome="1"							
	/clone="RAFL06-11-F13"							
	/ecotype="Columbia"							
	/note="This clone is in pBluescript"							
gene	1..1372							
CDS	/gene="Atlg44750; T12C22.2"							
	50..1189							
	/gene="Atlg44750; T12C22.2"							
	/codon_start=1							
	/product="Unknown protein"							
	/protein_id="AAL32860.1"							
	/db_xref="GI:17065412"							
	/translation="MAKEPVRVLVTGAAGNQEPILVKESVVGIPPTLLKLKQWQWV							
	LVSNIIFLIQQAQSVLLGRFYDEGNSKMATLVQTAAPFPIYPIPLLLPSSASV							
	ESSSSCKLYLVLYLVGLVGIAGDNLVSVGLLYLSASVSLICATOLAFNAFVS							
	FINAQKTFALLINLVLLSFAALIALNDADTPSGVSRSKIYGVFCTLAASLYSL							
	LLSLQFSKILKRETFPSVLEMQIYTSILVATCVIGLFPASGEWRTLHGEWGYHK							
	QASVLTWVTVQVCSVGVGLIPLVTSFSNVISTLSLATVPLAALVWFRDKM							
	SGVKIMAMLIAIWGFASYVYQNHIDDLKVRQARQAQGRVEPPC"							
ORIGIN								

RESULT 15	
TI2C22	
LOCUS	115421 bp DNA linear PLN 22-JUN-2000
DEFINITION	Sequence of BAC TI2C22 from Arabidopsis thaliana chromosome 1, complete sequence.
ACCESSION	AC020576
VERSION	AC020576.2 GI:6728952
KEYWORDS	HTG.
SOURCE	Arabidopsis thaliana (thale cress)
ORGANISM	Arabidopsis thaliana
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE	1 (bases 1 to 115421)
AUTHORS	Liu,S.; Vaysberg,M.; Sakano,H.; Lee,J.; Lenz,C.; Pham,P.; Toriumi,M.; Yu,G.; Chin,C.; Chlou,J.; Choi,E.; Chung,M.; Gonzalez,A.; Howng,B.; Liu,A.; Altafi,H.; Brookes,S.; Buehler,E.; Chao,Q.; Conn,L.; Conway,A.; Hansen,N.; Johnson-Hopson,C.; Khan,S.; Kim,C.; Lam,B.; Miranda,M.; Nguyen,M.; Palm,C.; Shinn,P.; Southwick,A.; Davis,R.; Ecker,J.; Federspiel,N. and Theologis,A.
	The sequence of BAC TI2C22 from Arabidopsis thaliana chromosome 1 Unpublished
TITLE	2 (bases 1 to 115421)
REFERENCE	Theologis,A.
AUTHORS	Direct Submission
TITLE	Submitted (05-JAN-2000) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA
JOURNAL	3 (bases 1 to 115421)
REFERENCE	Theologis,A.
AUTHORS	Direct Submission
TITLE	Submitted (09-FEB-2000) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA
JOURNAL	4 (bases 1 to 115421)
REFERENCE	Theologis,A.
AUTHORS	Direct Submission
TITLE	Submitted (22-JUN-2000) Plant Gene Expression Center, 800 Buchanan St., Albany, CA 94710, USA
JOURNAL	On Jan 21, 2000 this sequence version replaced gi:6671920.
COMMENT	The sequence of BAC TI2C22 from Arabidopsis thaliana chromosome 1.
FEATURES	Location/Qualifiers
Source	1..115421
	/organism="Arabidopsis thaliana"
	/mol_type="genomic DNA"
	/cultivar="Columbia"
	/db_xref="taxon:3702"
	/chromosome="1"
	/clone="TI2C22"
	185..1855
gene	/gene="TI2C22.1"
CDS	join(185..360,885..996,1263..1370,1697..1855)
	/gene="TI2C22.1"
	/codon_start=1
	/evidence=not experimental
	/product="TI2C22.1"
	/protein_id="AAF78257.1"
	/db_xref="GI:8655984"
	/translation="WTEGEINVLGIKEEGEINVLSTKEEGAKVFFGKIGRGSPRS CNIGALIGHLDIRKFPVLVVTFTIGVEDSPKLHNSIKFSGCKWRRRRYGVKKEA KTRDRGVEDATRKRGROVSTPASNGSASVLKIVGEKLMTSNFTVNPDSVSKGV WRNLTPPEGEIDIVAAATELYAM"
gene	5388..7868
CDS	/gene="TI2C22.2"
	join(5388..5430,6772..7868)
	/gene="TI2C22.2"
	/note="Contains similarity to purine permease from Arabidopsis thaliana gb AF078531. EST gb A1997301 comes from this gene."
	/codon_start=1
	/evidence=not experimental
	/protein_id="AAF78258.1"
	/db_xref="GI:8655985"
	/translation="WAKFPRVRLVLTGAAGNOBPILVKEESVVGITPPLLKLSQWQWV LVSVNIFFLIGQAASVLLGRFYDEGGNKMATLVQTAAFPILYIPLLLPSASV
gene	ESESSCSKYIVLIYVLLGVIIAGDNLMSVGLLYLSASTYSLICATQLAFNAVFSY PINAQFTEALINSVLLSFSAALANDADTPGVSRSRKYIVGFCVTLTAAASLYSL LLSLMQFSEFKILKRETFVSLEMOIYATSLVATCVSGLFASGFWRTILHGMEGYHK GOASVLTTLVMTAVTWQCSVGVGLIFLVTSLFSNVISTLSLATVPLAALVVRDKM SGVKIMAMLIATIGVFPASVYQNHIDDLKVRQARQAQRVEPPC"
	complement(9594..11127)
	/gene="TI2C22.3"
	complement(join(9594..9744,9836..9957,10759..11127))
	/gene="TI2C22.3"
	/note="Contains similarity to hypothetical protein F19B11.17 gi 4406763 from Arabidopsis thaliana BAC F19B11 gb AC006836."
	/codon_start=1
	/evidence=not experimental
	/protein_id="AAF78259.1"
	/db_xref="GI:8655986"
	/translation="MQFSRNSILRQLSRKEGWSRASKRWTSGDSTAFNDDTSGGGSY SMEGIVYSGDSTAARSKRVVVVDESSRKHAMWALTHTLTNGDILVTLHHVVSPL DEATPSLAQSLGSLCKACKPEVIALVIQGPKLATVLSQVKLEVSVLVLGQKKSAP LISLCGSPSRSEELVNRINCINGADCLITIGVRKQCKGVGGYLINTRQKNFWLLA"
gene	complement(20534..22109)
	/gene="TI2C22.4"
	complement(join(20534..20615,20695..20803,21229..21276,21381..21566,21643..21760,21843..22109))
	/gene="TI2C22.4"
	/note="Contains weak similarity to tail completion gi 5354213 from coliphage T4 gb AF158101. ESTs gb AA650799, gb AA041054, gb R29873, gb AA712908 come from this gene."
	/codon_start=1
	/evidence=not experimental
	/protein_id="AAF78260.1"
	/db_xref="GI:8655987"
	/translation="WTSHJAIEVTKVLEADVAVTAVETVHHHHHHQDENHSTNPI SDPRDEALRQENRRIRLTLESNLKLFETLAESAASFHDCPSDLYARLVMTVTSRD FLAKLENQALSNGTQNFPPFTEDDVKTVEVLIMDHQESFWMVLVTDMMVPSN VEQSADINHEIYVNEEHYIDAVAHLKATGFKLQNLKPEELQKLJLQVEVTSLSKV GKVDIIVHAGKWFYTLTWLAFGLYQARGVLKIAKGVHATSKVLRAL"
gene	22453..24681
	/gene="TI2C22.5"
	join(22453..22580,22790..22895,22987..23237,23328..23433,23515..23832,23910..23989,24078..24202,24407..24681)
	/gene="TI2C22.5"
	/note="Contains similarity to hypothetical protein T12G13.150 from Arabidopsis thaliana gi 7486671."
	/codon_start=1
	/evidence=not experimental
	/protein_id="AAF78261.1"
	/db_xref="GI:8655988"
	/translation="METAEMNHNGSKNLKNVDATEIEFKILAAALRSRVTYLRNEADC PTLVSRMLREDIGLEKCDLVYKSFVKEHLVKLEAGNNDTSENSQETEREDEI PTKVAFQSEHEHPMDAGENTSREAKDVKGKNETLQORDIKRALRKASVYKAN SETITMASLRLLBEDLKEKSDLPFKFINKELDEVLPDPAPKSTESIVKRVKK KVKSTPSKMVSSEYNSDSTEGNVNEEVAVKTMARKVLSPKPMGMRKSENGKV SRGKKAHTEIDSESDSGDSSEKSLQKTETATDVYGRVHKLVSCKSGMSVPPN IYKKAQAPOEKREAMLIELEQILAKEGLSSDPSALRIKEVKKRKNISRLGEGDTN NIVNRSRRSSTSAPPKPKVTAESSESEDEPESNEEESNEKAERGSGSSEVEE SENSEMENTLTSK"
gene	complement(25103..26276)
	/gene="TI2C22.6"
	complement(join(25103..25212,25310..25389,25470..25541,25625..25668,25739..25795,26013..26089,26174..26276))
	/gene="TI2C22.6"
	/note="Contains similarity to cation transport protein CHAC from Escherichia coli gi 2506988. ESTs gb AA605474, gb A1995104, gb R90162 come from this gene."
	/codon_start=1
	/evidence=not experimental
	/protein_id="AAF78262.1"
	/db_xref="GI:8655989"
	/translation="MMVFGYSLIWKTPFPDPESLPDFIKGYRRVHOGSTDHRTGPD FPGRTVLEAAHEEVLREBKQYDQKEYLDLFTQDSNASEPAGVMVYIAGPDKK SNNNYLGPAPLIEDIAKQIVKAKGSPGNRDYLFNLEEALAQLFKDKHVTDLANQVRH

gene	ILSSEELDIDATAATANNV complement(27042..29558) /gene="T12C22.7"	Qy	250	AAGACTAAACCATTACCAAAAAAATACTCTTCTCTTGACACTATCTTTAGTGTAT	309
CDS	complement(join(27042..27234,27331..27482,27571..27729, 27804..28044,28296..28412,29221..29283,29371..29558)) /gene="T12C22.7" /note="Contains similarity to Mtn21 gene from Medicago truncatula gb Y15293 and contains two integral membrane protein domains DUF6 of unknown function PF 00892. ESTs gb A1998702, gb 230851 come from this gene." /codon_start=1 /evidence=not experimental /protein_id="AAF78263.1" /db_xref="GI:8655990"	Db	7023	TCGGTAGTGTAGAGTCTTCGGAGAGTTCGTGTTCACTCAAGTACATTTGTTGATCTAT	7082
		Qy	310	ATTGGACTTCGGCTTGTGTTGCTGGACATTTGATTTTGTACTCATTTGGGCTACTTTAC	369
		Db	7083	GTTTTGCTCGGTGTGATCAATGCTGGAGATAATATGTTACTCTGTGGACTTTTGATC	7142
		Qy	370	CTTCTGTCTCAACTTCTTCTTGAATCTGCGTCGCAATTTGGCTTTTAACGCGCTTTC	429
		Db	7143	CTCTCTGCATCGAGTATTTCGCTCAATTTGGCGTACTCAGTTAGCTTTTCAACGCGGTGTC	7202
		Qy	430	TCCTTACTTCCTAAACTCACAAAAAATCACACCACTTATATCTCAATTCACCTTGTCTCTTA	489
		Db	7203	TCCTTATTTCAATGCTCAGAAGTTCACCTGCTTTGATTTCTCACTCCGTTGTTCTTGT	7262
		Qy	490	ACCATATCTTCTACATTTCTTGTATTCACATGAACCAAGATCTCCCTCTTCTACTTCA	549
		Db	7263	TCGTTCTCCGCTGCTTTGTAGTACTCTCAATGATGATCGCGATCTCTCTCTCGGT-----	7316
		Qy	550	AAGTCGACGCAAGTCCAAAGTATGATTTGGATACATCTGCGCGGTTCGGTAGCTCAGCT	609
		Db	7317	-----GTCTCCAGGCTAAAGTATATTTGTTGGGTTTGTGTACACTTGTCTGCGTCTGCT	7370
		Qy	610	GGTTATTTCTCTGCTGCTTCTTTAAACAGATTACGCGTTTCAAAAAGATTCTAAAGAAATAC	669
		Db	7371	CTCTATTTCTCTGTTGCTATCTTTATGTCAGTTCTGTTTCGAGAAGATTCTTGAAGAGGAG	7430
		Qy	670	ACATTTCAAGGCTATTTTGTAGACATGGCCACATATCCGCTCTATGTTAGTACTTGTGTAGTT	729
		Db	7431	ACGTTTCTCTGCTTCTTGAATGCAAACTCTACACTTCTTTAGTGGCGACTTGTGTTTCG	7490
		Qy	730	GTGTTAGACTTTTGTGAAGTGGTGGTGGAAAAGCTGAGTACAGAAATGGAGAGTTT	789
		Db	7491	GTTATAGGGCTTTTGTCTAGCGGGAAATGGAGAACGCTGCATGGGGAATGGAAAGGTTAT	7550
		Qy	790	CAACTAGGGAAAAGCTCATACATTTTGTATAAAACATCGGTTCAAACGATATCATGGCAAGCT	849
		Db	7551	CATTAAGGGCAAGCTCTTATGACTGACCTTGGTCTGGACAGCAGTTACTTGGCAAGTG	7610
		Qy	850	TGTTTGTATGGAAAGTGTGTTGATTTGATTTATCGAAGTTTTCATCGCTTTTTTCCAAATGTCATA	909
		Db	7611	TGTTCTGTTGGAGTCGTGGGTTTGTATTTCTGTTGACGTCGCTCTTCTCAAACGTCATT	7670
		Qy	910	AGCACTCTTTGTTTACAGTTTGGCTGTTCTTGTCTGTTCTTCTTCCGCTGATGAGATG	969
		Db	7671	AGTACGCTCTCTCTAGCTGTGACTCCACTAGCAGCTTTGGTTGTTTCCGCTGATAAAATG	7730
		Qy	970	AGTGGAAATCAAGTTGGTTGCAATGTTTTCGGCCATCTGGGGATTTGTTTCTTATGGTTAT	1029
		Db	7731	AGTGGTGTAAAGATTATGGCAATGCTGATCGCTATTTGGGGTTTCGCTTCTTATGTTTAC	7790
		Qy	1030	CAGATTATGTCATATGATGAAAGCCAGAGAACCAAGAGCTTCTCTCA	1079
		Db	7791	CAGAATCATATTGATGATCTGAAAGTAAGACAAGACGACGACAACAAGCTCA	7840
Search completed: November 1, 2004, 15:44:33					
Job time : 5081.95 secs					

gene	complement(27042..29558) /gene="T12C22.7"	Qy	250	AAGACTAAACCATTACCAAAAAAATACTCTTCTCTTGACACTATCTTTAGTGTAT	309
CDS	complement(join(27042..27234,27331..27482,27571..27729, 27804..28044,28296..28412,29221..29283,29371..29558)) /gene="T12C22.7" /note="Contains similarity to Mtn21 gene from Medicago truncatula gb Y15293 and contains two integral membrane protein domains DUF6 of unknown function PF 00892. ESTs gb A1998702, gb 230851 come from this gene." /codon_start=1 /evidence=not experimental /protein_id="AAF78263.1" /db_xref="GI:8655990"	Db	7023	TCGGTAGTGTAGAGTCTTCGGAGAGTTCGTGTTCACTCAAGTACATTTGTTGATCTAT	7082
		Qy	310	ATTGGACTTCGGCTTGTGTTGCTGGACATTTGATTTTGTACTCATTTGGGCTACTTTAC	369
		Db	7083	GTTTTGCTCGGTGTGATCAATGCTGGAGATAATATGTTACTCTGTGGACTTTTGATC	7142
		Qy	370	CTTCTGTCTCAACTTCTTCTTGAATCTGCGTCGCAATTTGGCTTTTAACGCGCTTTC	429
		Db	7143	CTCTCTGCATCGAGTATTTCGCTCAATTTGGCGTACTCAGTTAGCTTTTCAACGCGGTGTC	7202
		Qy	430	TCCTTACTTCCTAAACTCACAAAAAATCACACCACTTATATCTCAATTCACCTTGTCTCTTA	489
		Db	7203	TCCTTATTTCAATGCTCAGAAGTTCACCTGCTTTGATTTCTCACTCCGTTGTTCTTGT	7262
		Qy	490	ACCATATCTTCTACATTTCTTGTATTCACATGAACCAAGATCTCCCTCTTCTACTTCA	549
		Db	7263	TCGTTCTCCGCTGCTTTGTAGTACTCTCAATGATGATCGCGATCTCTCTCTCGGT-----	7316
		Qy	550	AAGTCGACGCAAGTCCAAAGTATGATTTGGATACATCTGCGCGGTTCGGTAGCTCAGCT	609
		Db	7317	-----GTCTCCAGGCTAAAGTATATTTGTTGGGTTTGTGTACACTTGTCTGCGTCTGCT	7370
		Qy	610	GGTTATTTCTCTGCTGCTTCTTTAAACAGATTACGCGTTTCAAAAAGATTCTAAAGAAATAC	669
		Db	7371	CTCTATTTCTCTGTTGCTATCTTTATGTCAGTTCTGTTTCGAGAAGATTCTTGAAGAGGAG	7430
		Qy	670	ACATTTCAAGGCTATTTTGTAGACATGGCCACATATCCGCTCTATGTTAGTACTTGTGTAGTT	729
		Db	7431	ACGTTTCTCTGCTTCTTGAATGCAAACTCTACACTTCTTTAGTGGCGACTTGTGTTTCG	7490
		Qy	730	GTGTTAGACTTTTGTGAAGTGGTGGTGGAAAAGCTGAGTACAGAAATGGAGAGTTT	789
		Db	7491	GTTATAGGGCTTTTGTCTAGCGGGAAATGGAGAACGCTGCATGGGGAATGGAAAGGTTAT	7550
		Qy	790	CAACTAGGGAAAAGCTCATACATTTTGTATAAAACATCGGTTCAAACGATATCATGGCAAGCT	849
		Db	7551	CATTAAGGGCAAGCTCTTATGACTGACCTTGGTCTGGACAGCAGTTACTTGGCAAGTG	7610
		Qy	850	TGTTTGTATGGAAAGTGTGTTGATTTGATTTATCGAAGTTTTCATCGCTTTTTTCCAAATGTCATA	909
		Db	7611	TGTTCTGTTGGAGTCGTGGGTTTGTATTTCTGTTGACGTCGCTCTTCTCAAACGTCATT	7670
		Qy	910	AGCACTCTTTGTTTACAGTTTGGCTGTTCTTGTCTGTTCTTCTTCCGCTGATGAGATG	969
		Db	7671	AGTACGCTCTCTCTAGCTGTGACTCCACTAGCAGCTTTGGTTGTTTCCGCTGATAAAATG	7730
		Qy	970	AGTGGAAATCAAGTTGGTTGCAATGTTTTCGGCCATCTGGGGATTTGTTTCTTATGGTTAT	1029
		Db	7731	AGTGGTGTAAAGATTATGGCAATGCTGATCGCTATTTGGGGTTTCGCTTCTTATGTTTAC	7790
		Qy	1030	CAGATTATGTCATATGATGAAAGCCAGAGAACCAAGAGCTTCTCTCA	1079
		Db	7791	CAGAATCATATTGATGATCTGAAAGTAAGACAAGACGACGACAACAAGCTCA	7840
Search completed: November 1, 2004, 15:44:33					
Job time : 5081.95 secs					

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 1, 2004, 05:48:03 ; Search time 96.7741 Seconds  
(without alignments)  
7704.726 Million cell updates/sec

Title: US-09-913-767-2

Perfect score: 1049

Sequence: 1 aagatgaagatgaagacagt.....actgattatgtacttaaaa 1049

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.\*

1: /cgn2\_6/prodata/1/ina/5A COMB.seq.\*

2: /cgn2\_6/prodata/1/ina/5B COMB.seq.\*

3: /cgn2\_6/prodata/1/ina/6A COMB.seq.\*

4: /cgn2\_6/prodata/1/ina/6B COMB.seq.\*

5: /cgn2\_6/prodata/1/ina/PCUS COMB.seq.\*

6: /cgn2\_6/prodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	73.6	7.0	7218	1	US-08-232-463-14
2	36.2	3.5	705	4	US-09-270-767-5061
3	36.2	3.5	705	4	US-09-270-767-20343
4	36.2	3.5	1827	4	US-09-270-767-1308
5	36.2	3.5	1827	4	US-09-270-767-16590
6	36	3.4	3969	4	US-09-248-796A-5440
7	35	3.3	282	4	US-09-313-294A-5650
8	35	3.3	7120	4	US-09-380-773-2
9	33.4	3.2	399	4	US-09-621-976-8976
10	33.4	3.2	3348	4	US-09-312-762A-2
11	33.4	3.2	14707	4	US-09-312-762A-3
12	33.4	3.2	23439	4	US-08-956-171E-38
13	33.4	3.2	23439	4	US-08-781-986A-38
14	33.2	3.2	289	3	US-09-007-005-17
15	33.2	3.2	289	3	US-09-244-796-17
16	33	3.1	5227	4	US-09-919-172-79
17	33	3.1	5228	4	US-09-919-039-216
18	33	3.1	15144	3	US-08-458-434A-6
19	32.6	3.1	3190	3	US-08-986-768-4
20	32	3.1	285	2	US-08-630-822A-85
21	32	3.1	285	3	US-09-005-069-85
22	32	3.1	285	3	US-09-171-156A-34
23	32	3.1	285	4	US-09-004-730A-34
24	32	3.1	285	4	US-08-981-799A-34
25	31.8	3.0	358	4	US-09-513-999C-24878
26	31.8	3.0	685	3	US-09-183-266A-16
27	31.8	3.0	996	4	US-09-543-681A-3267
28	31.8	3.0	1654	4	US-09-634-238-16
29	31.8	3.0	4177	3	US-09-023-082A-23
30	31.8	3.0	4177	4	US-09-248-998-23
31	31.8	3.0	4177	4	US-09-610-653-23
32	31.6	3.0	240	3	US-08-559-397A-10
33	31.6	3.0	276	4	US-09-248-796A-1906
34	31.6	3.0	381	4	US-09-248-796A-1906
35	31.6	3.0	1261	3	US-08-961-083-25
36	31.6	3.0	1261	4	US-09-536-784-25
37	31.6	3.0	1329	4	US-09-583-110-519
38	31.6	3.0	3096	4	US-09-614-221A-452
39	31.6	3.0	26385	4	US-08-961-527-3
40	31.4	3.0	1101	4	US-09-543-681A-3334
41	31.4	3.0	2243	1	US-07-995-657-1
42	31.4	3.0	2243	1	US-08-474-587-1
43	31.4	3.0	2871	4	US-09-489-847-111
44	31.4	3.0	5798	4	US-09-377-285B-33
45	31.2	3.0	711	4	US-09-107-532A-609

#### RESULT 1

US-08-232-463-14  
; Sequence 14, Application US/08232463  
; Patent No. 5670367

; GENERAL INFORMATION:

; APPLICANT: DORNER, F.

; APPLICANT: SCHEIFLINGER, F.

; APPLICANT: FALKNER, F. G.

; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS

; NUMBER OF SEQUENCES: 52

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Foley & Lardner

; STREET: 1800 Diagonal Road, Suite 500

; CITY: Alexandria

; STATE: VA

; COUNTRY: USA

; ZIP: 22313-0299

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent in Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/232,463

; FILING DATE:

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/07/935,313

; FILING DATE:

; APPLICATION NUMBER: EP 91 114 300.6

; FILING DATE: 26-AUG-1991

; ATTORNEY/AGENT INFORMATION:

; NAME: BENT, Stephen A.

; REGISTRATION NUMBER: 29,768

; REFERENCE/DOCKET NUMBER: 30472/114 IMMU

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (703)836-9300

; TELEFAX: (703)683-4109

; TELEX: 899149

; INFORMATION FOR SEQ ID NO: 14:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 7218 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; IMMEDIATE SOURCE:

; CLONE: pTZgpt-F1s

; US-08-232-463-14

Query Match

7.0%; Score 73.6; DB 1; Length 7218;

[illegible]

```

RESULT 2
US-09-270-767-5061/c
; Sequence 5061, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5061
; LENGTH: 705
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-5061

```

	Query Match	3.5%;	Score 36.2;	DB 4;	Length 705;
	Best Local Similarity	46.1%;	Pred. No. 0.099;		
	Matches 160;	Conservative 0;	Mismatches 183;	Indels 4;	Gaps 1;
Qy	119	TTCCAAAGCTTCCTTCAAAACGGTTGGTTGTGTCACATCATTTCTCCCTCTCTCTTATCTTT	178		
Db	593	TTCTTTTCTTTCCTCCCTCTCTTTTTCCTCTCTCTTTCTTTTTCCTCTCTCTCTCTCTTT	534		
Qy	179	TCCTCCGCGCTGCTGGTTGCTTGAAGAACAGACATCCATTTTTCTCATGAAC	238		
Db	533	TCTCTCATCTCTTTTCTTTTTTCCCTCTCTCTTTTCTCTCTCTCTTTTCTTTTCTTTTTC	474		
Qy	239	CTCCTCTCTTTATCCGCGCTATCGTTGTGTTTGTCTGGGATTTGCAAT----	294		
Db	473	CTCTCTCTTTTCTCTCTCTCTTTTCTTTTCTCTCTCTCTCTCTCTCTCTCTCTCTTC	414		
Qy	295	CTCTACTCTTACGGGTTAGCTTATACCTCGTTTCTACTCGGCTCTTGTGATCATCTCCGCG	354		
Db	413	TTTTTCTCGTTCCTCTCTCTTTTTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTTTT	354		
Qy	355	CAATTAGGCTTCACTGCTCTCTTTTGCAATTTTTTATGGTGAAGAAAAAGTTTCACACCTTTC	414		
Db	353	TCTCTGCTCTCTTTTCTTTTCT	294		

```

Oy      415 ACTATAAACCGCTATCGTTTTCCTCAGTGGTGGCCGTAAGTCCTTGC 461
          |||         |||         |||         |||         |||
Db      293 CTCCTCGCTTCTTTTTTTCCTTCTCTCTCTCTCTCTCTCTCTCTCTC 247

RESULT 3
US-09-270-767-20343/c
; Sequence 20343, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila m
; FILE REFERENCE: File reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20343
; LENGTH: 705
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-20343
```

	Query Match	3.5%; Score 36.2; DB 4; Length 705;
	Best Local Similarity 46.1%; Pred. No. 0.099;	
	Matches 160; Conservative 0; Mismatches 183; Indels	
Qy	119 TTCNAGCTCCCTCAAAACGGTTGGTGTGCCACTCATTTTCCTCCTGCTTC	
Dd		
Qy	593 TCTCTTTCTCTCTCCCTCTCTTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	
Dd		
Qy	179 TCCTCGCCGCTCGTGCTTGCTTGAAGAACAAGAACGACTCCATTTTCCC	
Dd		
Qy	533 TCTCTCATCT	
Dd		
Qy	239 CTCTCTCTTTATCGCGGTATCGTTGTGGTTGTCTGGGATTGTGACAA	
Dd		
Qy	473 CTCCTCTCTTTTCT	
Dd		
Qy	295 CTCTACTCTTACGGGTAGCTTATATCCCTGTTTCTACTGCGCTCTTTGATC	
Dd		
Qy	413 TTTTCTTCGFTC	
Dd		
Qy	355 CAATTAGGCTTCACTGCTCTCTTTGCAATTTTTATGGTGAAGCAAAAGTTCT	
Dd		
Qy	353 TCT	
Dd		
Qy	415 ACTATAAAGCTATCGTTTTGCTCACTGGTGGTGCCTAGTCTCTGCTGC 461	
Dd		
Qy	293 CTCCT	
Dd		

```

RESULT 4
US-09-270-767-1308/c
; Sequence 1308, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1308
; LENGTH: 1827
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-1308

Query Match          3.5%; Score 36.2; DB 4; Length 1827;
Best Local Similarity 49.0%; Pred. No. 0.19;

```



Matches	124;	Conservative	0;	Mismatches	128;	Indels	1;	Gaps	1;
Qy	126	CTTCCTTCAAACCGTGGTGTGCACACTCAATTTCTCCCTCTTCTCTTATCTTTCTCTCCG	185						
Db	1151	CTTTCCCTCTCCCTTTTCGTTTTTCCCTCTCTCTTTCTCTCTTTCTCTCTCTTTCTTTCTTCC	1092						
Qy	186	CCGTCGTCGTTCCTTCGAAGAACAGAAACGACATTTTCTCTCATGAACCTCTCTCT	245						
Db	1091	TCCTCTTTTTCCTATTATGTCCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1032						
Qy	246	CTTTTATCGCCGCTATCGTGTGTGTGCTCGTGGATTGACAATTAACCTTACTCTTA	305						
Db	1031	TTTTTTCTTTCCCTCTCTCTTCTCTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTTTT	972						
Qy	306	CGGGTTAGCTATATCCCTGTTTCTACTCGGTCTTTTGATCATCTCCGGGAATTAGCGTT	365						
Db	971	CTT-TCCCTCTCTCTTTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTTC	913						
Qy	366	CATGCTCTCTTTT	378						
Db	912	CTCTCTCTCTTT	900						

## RESULT 5

```

US-09-270-767-16590/C
; Sequence 16590, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16590
; LENGTH: 1827
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-16590

```

## RESULT 6

US-09-248-796A-5440/c  
; Sequence 5440, Application US/09248796A  
; Patent No. 6747137  
; GENERAL INFORMATION:  
; APPLICANT: Keith Weinstock et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN

```

; TITLE OF INVENTION:  FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 5440
; LENGTH: 3969
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-5440

```

## RESULT. T 7

```

US-09-313-294A-5650
; Sequence 5650, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 5650
; LENGTH: 282
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6476212 700350473H1
; NAME/KEY: unsure
; LOCATION: 2, 16, 149, 151, 241, 250, 274, 276
; OTHER INFORMATION: a, t, c, g, or other
US-09-313-294A-5650

```

### Query Match

3 38. Score 35. DB 4. Length 282.

Query Match 3 38. Score 35. DB 4

Query Match 3 38. Score 35. DB 4. Length 282.

### Query Match

Query Match 3 38. Score 35. DB 4. Length 282.

### Query Match

Query Match 3 38. Score 35. DB 4. Length 282.

```
Qy 872 TGATCTTGGCGTCA 887
Db 73 TACTGCTGGCGTCT 88

RESULT 8
US-09-380-773-2
; Sequence 2, Application US/09380773
; Patent No. 6759219
; GENERAL INFORMATION:
; APPLICANT: Hein, Silke
; APPLICANT: Soehling, Brigitte
; APPLICANT: Gottschalk, Gerhard
; APPLICANT: Steinbuechel, Alexander
; TITLE OF INVENTION: Methods for the Biosynthesis of Polyesters
; FILE REFERENCE: MORT136--- 118899.0136.NPUS00
; CURRENT APPLICATION NUMBER: US/09/380,773
; CURRENT FILING DATE: 1999-09-03
; PRIOR APPLICATION NUMBER: PCI/US97/03994
; PRIOR FILING DATE: 1997-03-03
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 7120
; TYPE: DNA
; ORGANISM: Clostridium kluyveri
US-09-380-773-2

Query Match 3.3%; Score 35; DB 4; Length 7120;
Best Local Similarity 53.2%; Pred. No. 1.2;
Matches 74; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

Qy 314 CTTATATCCCTGTTTCTACTGCGTCTTGATCATCTCCGCGCAATTAGGCTTCACTGTC 373
Db 1797 CTTATACTAAGGCATCTACTCGCGGATTTATCTGTACTAATGCAGTATTTACCATA 1856

Qy 374 TCTTTGGATTTTATGTTGTAAGCAAGTTTCACACCTTTCACATATAACGCTATCGTT 433
Db 1857 CTTTTCGATATTTTCATATATAAAGAAAGATAAAGGATTAACCATAGTTTCTATTATTG 1916

Qy 434 TGCTCACTGGTGGTCCCT 452
Db 1917 TTTCACTGATGGTGACT 1935

RESULT 9
US-09-621-976-8976/c
; Sequence 8976, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 8976
; LENGTH: 399
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-8976

Query Match 3.2%; Score 33.4; DB 4; Length 399;
Best Local Similarity 13.3%; Pred. No. 0.6;
Matches 27; Conservative 102; Mismatches 73; Indels 1; Gaps 1;

Qy 699 GGTATACAGAGAGAAGAGATTTTAAAGCTTGGAGAGTCTTTGTACTATGTGGTAT 758
Db 260 GSYVRMAGYRSRWSYTSAMWRKKMTCKGRSSWGSRSSTGYVAMMYKKSMTSRKMY 201

Qy 759 TGTGTTACGGCCATAATCTGGCAAGCATTTTTTGTGGAGCTATTGGGTGATCTTCTG 818
Db 200 YKRRKKRRKCTSTRTCTYRGSTYKCAYYTKGRKKTWTYYYYSKYSMSMKKTWRMK 141

Qy 819 TGCATGCTCTCTGCTCTCTGGAATATGATGTCAGTGTCTCTCCGGTGACGGTGATCTT 878
Db 140 TAYWTGKWKWT-RTKWTCTMCWCTTYWAGTMYRYRYRYWYAKRAKWSKRCWTSTT 82

Qy 879 GGCGCTCAITTGCTCCAGGAGA 901
Db 81 CYCMKYWAKKWSYVWSMSMMKW 59

RESULT 10
US-09-312-762A-2
; Sequence 2, Application US/09312762A
; Patent No. 6552177
; GENERAL INFORMATION:
; APPLICANT: MIA HOROWITZ ET AL.
; TITLE OF INVENTION: BH DOMAIN CONTAINING GENES AND PROTEINS
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Mark M. Friedman c/o Anthony Castorina
; STREET: 2001 Jefferson Davis Highway, Suite 207
; CITY: Arlington
; STATE: Virginia
; COUNTRY: United States of America
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 1.44 megabyte, 3.5" microdisk
; COMPUTER: Twinhead; Slimnote-890TX
; OPERATING SYSTEM: MS DOS version 6.2,
; OPERATING SYSTEM: Windows version 3.11
; SOFTWARE: Word for Windows version 2.0 converted to
; SOFTWARE: an ASCII file
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/312,762A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/026,898
; FILING DATE: 20 FEB 1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Friedman, Mark M.
; REGISTRATION NUMBER: 33,883
; REFERENCE/DOCKET NUMBER: 916/10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 972-3-5625553
; TELEFAX: 972-3-5625554
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3348
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-09-312-762A-2

Query Match 3.2%; Score 33.4; DB 4; Length 3348;
Best Local Similarity 49.7%; Pred. No. 2.6;
Matches 85; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

Qy 832 GTCTCTGGAATTATGTCAGTGTCTGCTTCCGGTGACGGTGATCTTTGGCGCGTATTC 891
Db 1226 GTGCGATATGCTGGCCAACGATATAGCTCGGTGATGTTGGTGGTGGAGGAG 1285

Qy 892 TTCCAGAGAGATTTCAGCGGGGAAGGTGCTGCTTGGCTCTCTCCCTCTGGGATCA 951
Db 1286 TCCTGTATGCCCTCACAGGCTGTGAAGGGTGGTGTCTTTTATGGCACCATGAATGGGCC 1345

Qy 952 GTCTCTTATTCTATGTCAGAGGTTAAATCCGAGGAGAAGACTAAGGCTCAG 1002
```

Db 1346 TTTGGGCGATGGCTACGCGGAGGGGGCTGGCGAGGGGCAATTGATGATGTTGAG 1396

## RESULT 11

US-09-312-762A-3

; Sequence 3, Application US/09312762A

; Patent No. 6552177

; GENERAL INFORMATION:

; APPLICANT: MIA HOROWITZ ET AL.

; TITLE OF INVENTION: EH DOMAIN CONTAINING GENES AND PROTEINS

; NUMBER OF SEQUENCES: 27

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Mark M. Friedman c/o Anthony Castorina

; STREET: 2001 Jefferson Davis Highway, Suite 207

; CITY: Arlington

; STATE: Virginia

; COUNTRY: United States of America

; ZIP: 22202

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 1.44 megabyte, 3.5" microdisk

; COMPUTER: Twinhead\* Slinnote-890TX

; OPERATING SYSTEM: MS DOS version 6.2,

; SOFTWARE: Word for Windows version 3.11

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/312,762A

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 09/026,898

; FILING DATE: 20 FEB 1998

; ATTORNEY/AGENT INFORMATION:

; NAME: Friedmam, Mark M.

; REGISTRATION NUMBER: 33,883

; REFERENCE/DOCKET NUMBER: 916/10

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 972-3-5625553

; TELEFAX: 972-3-5625554

; TELEX:

; INFORMATION FOR SEQ ID NO: 3:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 14707

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

US-09-312-762A-3

Query Match 3.2%; Score 33.4; DB 4; Length 14707;

Best Local Similarity 49.7%; Pred. No. 7;

Matches 85; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

QY 832 GTCTCTGGAAATATGGTCAGTCTCTGCTTCGGGTGACGGTGATCTTGGCGGTCAATTGCG 891

Db 8947 GTGGATGATATGTCGCCCAACATATACTCGGCTGATGGTGATGGTGGCCAGGAGG 9006

QY 892 TTCCAGGAGAAGTTTCAGCGGGGAAAGGTGTCGGTCTCTCCCTCTGGGGATCA 951

Db 9007 TCCTGATGCCCTCAGAGCTGTGAGGTGTGCTTTTGATGGCACCAGTAATGGGCC 9066

QY 952 GTCTCTATTCTTATGGACAGGTTAAATCCGAGGAGAACTAAGGCTCAG 1002

Db 9067 TTTGGCATGGCTACGGCGAGGGGCTGGCGAGGGCAATTGATGATGTTGAG 9117

RESULT 12

US-08-956-171E-38

; Sequence 38, Application US/08956171E

; Patent No. 6591114

; GENERAL INFORMATION:

; APPLICANT: Charles Kunsch

; Gil H. Choi

; Patrick S. Dillon

; Craig A. Rosen

; Steven C. Barash

; Michael R. Fannou

; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences

; NUMBER OF SEQUENCES: 5256

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Human Genome Sciences, Inc.

; STREET: 9410 Key West Avenue

; CITY: Rockville

; STATE: Maryland

; COUNTRY: USA

; ZIP: 20850

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage

; COMPUTER: HP Vectra 486/33

; OPERATING SYSTEM: MSDOS version 6.2

; SOFTWARE: ASCII Text

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/956,171E

; FILING DATE: 20-Oct-1997

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 60/009,861

; FILING DATE: January 5, 1996

; APPLICATION NUMBER: 08/781,986

; FILING DATE: January 3, 1997

; ATTORNEY/AGENT INFORMATION:

; NAME: Mark J. Hyman

; REGISTRATION NUMBER: 46,789

; REFERENCE/DOCKET NUMBER: PB248P1

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (240) 314-1224

; TELEFAX: (301) 309-8439

; INFORMATION FOR SEQ ID NO: 38:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 23439 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; SEQUENCE DESCRIPTION: SEQ ID NO: 38:

US-08-956-171E-38

Query Match 3.2%; Score 33.4; DB 4; Length 23439;

Best Local Similarity 60.4%; Pred. No. 9.6;

Matches 55; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY 688 GCGGATTTCAAGGTGATAGGAGGAGACCAAGAGATTTTAAGCTTGGAGAGCTCTTTGTAC 747

Db 13437 GGTGATTATATGTTAAGAGGACAAAGAGAAAGATAGTATTAGAAAGTATTCAATA 13496

QY 748 TATGTGGTGATTGTGTTCCACGCCATAATCT 778

Db 13497 GCGTGGTGTCAGTGTAGCGGCTACAATGT 13527

RESULT 13

US-08-781-986A-38

; Sequence 38, Application US/08781986A

; Patent No. 6737248

; GENERAL INFORMATION:

; APPLICANT: Charles Kunsch

; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences

; NUMBER OF SEQUENCES: 5255

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Human Genome Sciences, Inc.

; STREET: 9410 Key West Avenue

; CITY: Rockville

; STATE: Maryland

; COUNTRY: USA

; ZIP: 20850

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage

; COMPUTER: HP Vectra 486/33



**DQ**      380 CA TTTT TATG TGTAAGCAAAAGTTCACACC TT 413  
             :  
**DB**      44 AYATTYTGYYAYAYATYA GYTAYAYATTY 11

Search completed: November 1, 2004, 22:40:40  
Job time : 104.774 secs

**THIS PAGE BLANK (USPTO)**

	Query Match	32.2%;	Score 337.6;	DB 16;	Length 1233;
	Best Local Similarity	61.5%;	Pred. No. 8.2e-97;		
	Matches 563;	Conservative	0;	Mismatches 344;	Indels 9; Gaps 1
y	80	TGCGTCTCTACTTCAAATAATGTGGCGAAGATCTGGTTTCCAAAGCTTCTTCAACCG	139		
b	2	TGCGTCTCTACTTCATCCACGGTGCGCAAACGAATCTGGCTCTCTAGCTTCTCTTGAACATG	61		
y	140	TTGGTTGGCCAACATATTTCCTCCCTCTCTCTCTATCTTTCTCCGCCGTGCTGTTGCC	199		
b	62	CAGCCAAATCGGCTCGAGTCCTCCCAATAATCATTTCTTAGATTCATAAACGACGTCGTC	121		

```
QY 200 TTGAAGAACAAAGAACGACTCATTTTCTCATGAAACCTCCTCTCTTTATCGCGCTA 259
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
122 ATCGCT-----CTTCGCGCTAAATCTCTATCAAGCCCCCTCTCTCTCGCCCTCG 172
QY 260 TCGTGTGTTGTTGTCGTCGGGATTTGACAAATTAACCTCTACTCTTACGGGTAGCTTATA 319
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
173 CCTCATCGGACTCTCAACCGGCTCGAGGACTACCTCTACGCTTGGGCGTGCTCGCC 232
QY 320 TCCCTGTTTCTACTCGGCTTTTGATCATCTCGGGCAATTAGGCTTCACTGCTCTCTTTG 379
Db 233 TTCCGGTCTCACCTTCTCTCTAATCCAAAGCTCCACCTCGGCTTACCGCCGCTCTTCG 292
QY 380 CATTTTATGTTGTAAGCAAAAGTTTCAACCTTTCACTATATAACGCTATCGTTTGCTCA 439
Db 293 CTTTCTCTCTCGTCGCCACAGTTTCAAGCCCTACTCCGTCACCTCCGTCGCTGCTTCTCA 352
QY 440 CTGTTGTTGTCGCTAGTCTTCCCTTAACTCTGATAGTGACAAGCTTGCAAAACGAGAC 499
Db 353 CGTCGCCGCTGTGTTCTTGCTCTGCTCTCGTTCCAGGGGACCGCCCGGTCAGTGA 412
QY 500 ACAAGGAATATGTTGTTGGGTTCAATGATCACTCTTTGGTGACGCTTCTCTATGGGTTTA 559
Db 413 GTCGCCAGTACGTTGATTTGTTTATGATGACTTTCGCGCTCGCGGCTGTATGGGTTTCG 472
QY 560 TATTGCCACTTGTGAGCTTTCTTACAAAGAAATCTGTTGACGGAATCAAGTATACGCTCG 619
Db 473 TTTTGCCCTTGATGAGTTGTTGTATGATGTTTACAAAAAGAGCAGGACGATACAGTACTCTG 532
QY 620 CGCTCGAGTTCCAGATGTTTATGCTTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 679
Db 533 TCATGGAGATTCAAGCTTCTTGTGCTTCTTGTGCTTCTTGTGCTTCTTGTGCTTCTGCT 592
QY 680 TAGCCGCTGGGATTTCAAGGTATGATGAGAGAGAGAGAGAGATTTTAAGCTTGGAGAT 739
Db 593 TAATCAATAATGACTTCAAGGTGATTCGCGAGAGCAAGAGATTTTAAGCTTGGGAAA 652
QY 740 CTTTGTACTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 799
Db 653 CAAGTACTAGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 712
QY 800 CTATTGGGTTGATCTTCTGTCATGCTCTGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCT 859
Db 713 CAATAGGGGTTATCTTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 772
QY 860 TTCCGGTACCGGTGATCTTGGCGCTATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 919
Db 773 TTCCAGTGACGGAAGTTTGGCTGTTATTTATATACAAAGAGAGCTTTTCATGCAGAGAA 832
QY 920 GTGTGCTTTGCTCTCTCCCTCTGGGATCAGTCTCTTATTTCTATGACAGGTTAAAT 979
Db 833 GGGTTGCTTTGCTCTCTCTTTTGGGGTTTGTCTCTATTTCTATGGAGATATAAAC 892
QY 980 CCGAGGAGAGACTAA 995
Db 893 AAGACAGGAAAAAGAA 908
```

## RESULT 2

US-10-425-114-10516  
; Sequence 10516, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:

; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(5313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 10516  
; LENGTH: 1146  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: 700941791\_FLI  
US-10-425-114-10516

Query Match 30.1%; Score 316.2; DB 16; Length 1146;  
Best Local Similarity 63.2%; Pred. No. 5.6e-90;  
Matches 486; Conservative 0; Mismatches 283; Indels 0; Gaps 0;

```
QY 227 TCCTCATGAACCTCCTCTCTTTATCGCCGCTATCGTTGTGGTTGCTCGTGGGATTG 286
Db 61 TCTCTATCAAGCCCTCTCTCTCTCGCCCTCGGCTCATCGGACTCTCACCGCCCTCG 120
QY 287 ACAATTACCTCTACTCTTACGGGTTAGCTTATATATCCCTGTTTCTACTCGGCTCTTTGATCA 346
Db 121 ACGNCTACCTCTACGCTCGGCTGGCTCGCTCGGCTTCCGCTCTCCACTTTCTCTAATCC 180
QY 347 TCTCCGCGCAATTAGGCTTCACTGCTCTCTTTGCAATTTTATGTTGAAGCAAAAGTTCA 406
Db 181 AAGCCTCCCACTCGCCTTCAACGCGCTCTTGGCTTCTCTCTCTCTCTCTCTCTCTCT 240
QY 407 CACCTTTCACTATAAAGCTATCGTTTGTCTCACTGGTGGTGGCTAGTCTCTTGGCCTTA 466
Db 241 CGCCCTACTCCGTAACCTCGTCTGCTCTCTACCCGCTCGGCTGCTGGCTCTGCTCG 300
QY 467 ACTCTGATAGTACAAAGCTTGCAAAACGAGACACACAAAGGAATATGTTGTTGGGTTCA 526
Db 301 GTTCCAGCGGGACCGCCCGCGGTGAGTCCGAGTCCAGTACGTCATGTTGGTTTGTTA 360
QY 527 TGACTCTTGTGTGAGCTCTTCTCTATGGGTTTATATATGTTGCTTGTGCTGCTTCTTACA 586
Db 361 TGATACTTGGGCTGCGCGCTGTATGGGTTGCTTTTGGCTTGTATGAGTGGTGTGATCA 420
QY 587 AGAAATCTGTCAGGCAATCACTATACGCTCGGCTCGGCTCGAGTCCAGATGCTCTTATGCT 646
Db 421 AAAAGACAGGACGCGCTATCACGTACTCTCTGGTCATGGAGATTCAGCTTGTCTTGTGCT 480
QY 647 TTGCTGCACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 706
Db 481 TCTTTGTACTCTTATCTGCAACCGTTGGAATGATATCAATATGACTTCAAGGTGATTC 540
QY 707 CAGAGAAAGCAAGAGATTTTAAAGCTTGGAGAGCTCTTTGTAATGTAATGTTGTTGTTCA 766
Db 541 CGCAGAGAAAGCAAGAGATTTTAAAGCTTGGGAAAACAAAAGTACTACGTTGTGTGTTGGA 600
QY 767 CGCCATAATCTGCAAGCATTTTGTGGGAGCTATTTGGGTTGATCTTCTGTGCTATCGT 826
Db 601 GTGCAATAATGTCGAGCTTTTCTTTTGGAGCAATAGGGGTTATCTTTTGTGCTCTGCT 660
QY 827 CTCTGGTCTCTGGAATTAATGCTCAGTCTCTGCTTCCGCTGACGGTGATCTTGGCCGCTCA 886
Db 661 CTTTGTGTCGGTATTAATAATTCCTGCTTTCTTCCAGTGACGGAAGTTTGGCTGTTA 720
QY 887 TTTGCTTCCAGGAAAGTTTCAAGCGGGGAAAGGTTGCTGCTTGTGGCTCTCTCCCTCTGGG 946
Db 721 TTGTATACAAAGAGAGCTTTCATGCAGAGAAAGGGGTTGCTTGTGGTGTCTCTCTCTTTGG 780
QY 947 GATCAGTCTCTTATTTCTATGAGCAGGTTAAATCCGAGGAGAGACTAA 995
Db 781 GGTTTGTGCTCTATTTCTATGAGAGATATAAACAGACAGGAAAGAA 829
```

## RESULT 3

US-10-424-599-73501  
; Sequence 73501, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K



; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 73501  
; LENGTH: 994  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_37387C.1  
US-10-424-599-73501

Query Match 30.1%; Score 316; DB 16; Length 994;  
Best Local Similarity 61.6%; Pred. No. 6e-90;  
Matches 541; Conservative 0; Mismatches 330; Indels 7; Gaps 2;  
QY 10 ATGAAGACAGTTCTTGTAACTATAAATCTTGGCCATTGGAACTGTGGAGC 69  
DB 118 ATGAAGCGCTTCTTCAGCAAACTGTCTTACTCACCATCGGCACTCCGGTGG 177  
QY 70 CCTCTAATGATGGTCTCTACTTCCAAATGTTGGCGAAAGATCTGTTTCCAACTTC 129  
DB 178 CCCCTCGTATCGTCTCTACTTCTCCACGGCGGCCACCGGCTCTGGCTCTCCAGCTTC 237  
QY 130 CTTCAACCGTGTGTGTCACATCATTTTCTCCCTCTCTCTTATCTTCTCCGCCGT 189  
DB 238 CTCGAACCGTGGCTTCCCTCTCATGCTCTCCCTCCCTCGCGCTCTCTACCTCCGCCGA 297  
QY 190 CGTCG-----TTGCCCTTGAAGAACAAAGAACGACTCCACTTTTCTCTATGAACCTCT 243  
DB 298 CGTCGACCGCTCCGCGCGCGGACCGCAACGAATTAATCTCAATGAAGCTCTCT 357  
QY 244 CTCCTTATCGCGCTATCGTTGTTGTTGCTCGTGGGATTTGACAAATPACTCTACTCT 303  
DB 358 CTCCTCGCGCTCCCACTTTCATCGGAATCTCACGGGCTCGAGACTACCTCTACGCC 417  
QY 304 TAGGGTTAGCTTATATCTCTTCTACTGCGTCTTTGATCATCTCCGCGCAATTAGGC 363  
DB 418 TAGGGGTGGGAGGCTTCGGCTCTCACTTCGGCCCTCATATCGCAACGCAACTCGGC 477  
QY 364 TTCACCTGCTCTCTTGCATTTTATGTTGAAGCAAAAGTTCACACCTTTTCACTATAAC 423  
DB 478 TTCACCGGCTTCTTCGGCTTCTCTCGTAGCAGAGATTACGGGCTACTCCGTAAC 537  
QY 424 GCTATCGTTTGTCTCACTGGTGGTGGCGPAGTCTTTCGCCCTTAACTCTGATAGTACAAG 483  
DB 538 GCGCTGTTTGTCTCACTGTGCGCGCGCGGTTTGGGCTTTCACACGAGGAGACCGT 597  
QY 484 CTTGCAACGACACACAGAAGAAATATGTTGGTTTATCATGACTCTTGGTGCAGCT 543  
DB 598 CCCCCGCGAGTCCGTTAAGAAATATGTTATGGGCTTTGTGATGACATGATCGCTGCG 657  
QY 544 CTTCTCTATGGGTTTATATGTCACCTGTGAGCTTTTCTTCAAGAAATCTGGTCAGGA 603  
DB 658 GCATGTATGATTCATTTTACCTTGGTGGAGTTGGGTACAAAATATCAACAGCCT 717  
QY 604 ATCAGTATACGCTCGCGCTCGAGTTCAGATGGTCTTATGCTTTGCTGCACTTGTGTC 663  
DB 718 CTTACTTACTCTCTTGTATGAGATTCAAGTTCTGTTATGTGCTTCTCGGCCACTCTCTT 777  
QY 664 TGCCTCGTGGGATGCTAGCCCTGCGGATTTCAAGGTGATAGCAGGAGAACAGAGAT 723  
DB 778 TGCCTCTTGGAAATGATCATCAACAATGACTTTAAGGTGATTCGAGGGAAGCCAAAAA 837  
QY 724 TTTAAGCTTGGAGAGTCTTTGTACTATGTTGGTATGTTGTTTCCAGCGCCATATCTGGCAA 783  
DB 838 TTTGAGCAGGAGAGGAAGTTACTATGCTGTTTGGTGGGAGTGCATAATATTATGGCAG 897  
QY 784 GCATTTTGTGGGAGCTATTGGGTTGATCTTCTGTGCATCGTCTCTGCTCTGGAATT 843

DB 898 GCTTTTCTTGGGCGCATTTGGGTTATATTTTGGGCTGCTCTGATCT-AGGATTT 956  
QY 844 ATGGTCAGTCTCTCTTCCGGTGACGGTGATCTTGGC 881  
DB 957 TTGATTGGGTTGTTGCTATACCCGTAACGGAAGTGTGGC 994  
RESULT 4  
US-10-437-963-57539  
; Sequence 57539, Application US/10437963  
; Publication No. US2004012343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 57539  
; LENGTH: 1173  
; TYPE: DNA  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_59340C.1  
US-10-437-963-57539

Query Match 22.1%; Score 231.4; DB 17; Length 1173;  
Best Local Similarity 54.1%; Pred. No. 8.1e-63;  
Matches 540; Conservative 0; Mismatches 426; Indels 33; Gaps 2;  
QY 9 GATGAAGACAGTCTTGTAACTATAAATCTGTTGGCCATTGGAACTGTGGAGC 68  
DB 102 GTTCAGGAGCGGCTCTCGTCACTTCGTCCTCATGTTGGTGGCTCGGCGTGGCG 161  
QY 69 CCCTCTAATGATGCTCTCTACTTCCAAATATGTTGGCGAAAGATCTCGTTTCCAAAGCTT 128  
DB 162 GCGGCTCTCTCTCGGCTTACTTCTCCGCGGCGGCAACCGCAAGTGGCTCTCAGCCT 221  
QY 129 CTTCAAAACGTTGGTGTCCACTCATTTTCTTCCCTCTTCTCTTATCTTCTCCCGCG 188  
DB 222 TCTCAGACCGCGGCTGGCGCTGCTCTCGCGCGCTCTGCTCTCTCTCTCTCTCAG 281  
QY 189 TCGTGTGCTTGAAGAACAGAA-----CGACTCCATT 224  
DB 282 CCGCGCGCGCGAGGTTGAGGACGACGCGCTGGCGCGCGCGCCACCGCT 341  
QY 225 TTTCCTCATGAACCTCTCTTATCGCGCTATCGTTGTTGTTGCTGCTGGGATT 284  
DB 342 GTTCTCATGACGCAACGCTCTCGTGGCGCTCGCGCTCGTGGGCTCATGACCGCGT 401  
QY 285 TGACAAATPACTCTACTCTTACGGGTTAGCTTATATCCCTGTTTCTACTGGGCTTTGAT 344  
DB 402 CGAGACCTCTCTACGCTACGCTACGCGCTACCTCCCGGTGTCCACTCTCTCATCT 461  
QY 345 CATCTCCGCGCAATTAGGCTTCACTGCTCTTTCGATTTTTTATGTTGAAGCAAAAGTT 404  
DB 462 CATCTCAGCAGCTGGCTTTCAGCGCGGCTTTCGCTGCTGCTGCTGCGCGAGCGGT 521  
QY 405 CACACTTCTACTATAAACGCTATCGTTTTCCTCACTGCTGCGCGCTGCTGCTGCT 464  
DB 522 CACGGGTTCTCGGTGAACGCGCTGCTGCTCTAGCGCTCGCGCGCGGATGCTGGGAT 581  
QY 465 TAACCTCTGATAGTGACAGCTTGCACACGAGACACACAGGAATATGTTGTTGGTTCAT 524

582	GAACGCGGCGGGAGCCGCGCGGGGGTGTCTCGCGCGCGAGTACTGCGCGGGTTCG	641
525	CATGACTCTTTGGTCAGAGCTCTTCTCTATGGGTTTATATTGGCACTTGTTCGAGCTTTCTTA	584
642	CATGACGCTCGCGCGCGGGCGGTGTACGGGCTCGTGTCTCCCGTTCATGGAGCTCAGCCA	701
585	CAAGAA-----ATCTGGTCAGCGAATCAGTATACGCTCGCGCTCGAGTTCACAGT	635
702	GGCGCACCCAGCGCGCGCGCGCGCGTCAAGTACACGCTCGCTCATGGAGATTCGACGT	761
636	GGTCTTATGCTTTTGTCTGCCACTTGTGTCTGCCTCGTGGGATGCTAGCGCTGGCGATTT	695
762	CGTCATCGGCTTCGTGCGCACGGCCCTCAGCGCGCGTGGCATGCTCGTCAACACGATTT	821
696	CAAGGTGATAGCAGGAGAGCAAGAGATTTTAAAGCTTGGAGAGTCTTTGTACTATGTGGT	755
822	CCACGCAATCCACAGGAGAGCCCATGAATTCGGTCTGGGGCAAGCGGGCTACTACCTGCT	881
756	GATTGTGTTTCACGGCCATAATCTGGCAAGCATTTTTTGTGGAGCTATTGGGTTGATCTT	815
882	CCTGGCCGGATCGCGCGCCATGTACAGTGTCTTCTTCTCGGCACGATCGCGGCCATCTTT	941
816	CTGTGCAATCGTCTCTGTGCTCTCGAAATTATGGTCAGTGTCTGTCTCGGTGACGGTGAAT	875
942	CTACGGCTCGCGCGTGTCTCGCGCGGTCAATCATACCGGTGTCTCATCCGGGTACCCGAGT	1001
876	CTTGGCCGTTCATTTGCTTCCAGGAGAAAGTTTTCAGCGCGGGAAGAGGTGTCGCTTTGGCTCT	935
1002	GCTCGCCGTCAATGTTCTTCCACGAGCCATTCACGGCACAAAGGGCGTCCGCGTCTGCTCT	1061
936	CTCCCTCTGGGATCAGTCTCTTTATTTCTATGGACAGGT	974
1062	CTCGCTATGGGCTTCGTCTCTCTACTTCTACGGCGAGGT	1100

## RESULT 5

US-09-938-842A-2212  
; Sequence 2212, Application US/09938842A  
; Patent No. US20020160378A1  
; GENERAL INFORMATION:  
; APPLICANT: Harper, Jeff  
; APPLICANT: Krepes, Joel  
; APPLICANT: Wang, Xun  
; APPLICANT: Zhu, Tong  
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
; TITLE OF INVENTION: SAME, AND METHODS OF USE  
; FILE REFERENCE: SCRI1300-3  
; CURRENT APPLICATION NUMBER: US/09/938,842A  
; CURRENT FILING DATE: 2001-08-24  
; PRIOR APPLICATION NUMBER: US 60/227,866  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: US 60/264,647  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/300,111  
; PRIOR FILING DATE: 2001-06-22  
; NUMBER OF SEQ ID NOS: 5379  
; SEQ ID NO 2212  
; LENGTH: 3387  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
; US-09-938-842A-2212

	Query Match	16.0%	Score 167.6;	DB 9;	Length 3387;
	Best Local Similarity	48.6%;	Pred. No. 4e-42;		
	Matches 492;	Conservative 0;	Mismatches 514;	Indels 6;	Gaps 1;
Qy	1	AAGATGAAGATGAAGACAGTCTTTGTTAATCATAAACGTATATTTCTTTGGCCATTGGAAC	60		
	2353	AAGAACTGTAAGAGGTGGCTCCGTGCTCCATATACGCAATCTTTGTTCATCTTCTGCGAA	2412		
Db					
Qy	61	TGTGGAGGCCCTCTTAATGATCCGCTCTTACTTCCAAAATGTTGGCGAAAGGATCTGGTTT	120		
Db	2413	CCACTTGTCAAGTCTTGGGTGAGCTGTACTATCAAAATGAGCGGAAAGACACATATGTG	2472		

## RESULT 6

US-09-938-842A-2212  
 ; Sequence 2212, Application US/09938842A  
 ; Publication No. US20040009476A9  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Harper, Jeff  
 ; APPLICANT: Kreps, Joel  
 ; APPLICANT: Wang, xun  
 ; APPLICANT: Zhu, Tong  
 ; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING SAME, AND METHODS OF USE

```
; FILE REFERENCE: SRR1300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 2212
; LENGTH: 3387
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-2212

Query Match      16.0%; Score 167.6; DB 11; Length 3387;
Best Local Similarity 48.8%; Pred. No. 4e-42;
Matches 492; Conservative 0; Mismatches 514; Indels 6; Gaps 1;

QY 1 AAGATGAAGATGAAGACAGTCTTGTAAATCATAAACTGTATATCTTGGCCATTCGAAAC 60
   |||||
Db 2353 AAGACTGTAAAGGTGGCTCGGTCTCCATATAGCCTCTTGTCTATCTTCGCCAA 2412

QY 61 TGTGAGGCCCTCTAATGATCGCTCTACTTCCAAAATGGTGGGAAAGAGATCTGGTTT 120
   |||||
Db 2413 CCACCTTGTACAGTCTCGGTAGACTGTACTATGAAATGGAGGAAAGACATATGTG 2472

QY 121 CCAAGCTTCCTCAACCGTGGTGTGTCACCTCATTTTCTCCCTCTCTCTATCTTTC 180
   |||||
Db 2473 GTAACACTTCTCAACTCATTCGCTTCCCTGTAAGTCTCTTCGCTCTCTTCTCGA 2532

QY 181 CTCGCGCGCTGCTGCTTGAAGAAACAAAGAAAGACTCCATTTTCTCATGAACCT 240
   |||||
Db 2533 ATCAGGCAACCAATCAACAGATCAAAATTCAGTCAGTCCCTCTCTTACACCCCT 2592

QY 241 CTTCTCTTTATCCCGCTATCGTGTGTTGCTCGTGGGATTTGACAAATTAACCTCTAC 300
   |||||
Db 2593 GCATCGGTTTAC-----TTGTGCACTGGACTGCTAGTGTCTGCTTATTTGCT 2646

QY 301 TCTTACGGTGTAGCTTATATCCCTGTTTCTACTGGCTTTTGATCATCTCCGCGCAATTA 360
   |||||
Db 2647 GCAGTGGGTGCTTTACTTACCACTCTACTTTCTCCCTCATCTTGCCCTCACAGTTG 2706

QY 361 GCGTTCACGTCTCTTTGCAATTTTATGGTGAAGCAAAAGTTCACACCTTTTCACTATA 420
   |||||
Db 2707 GCGTTCACGTCTTTTCTATATTTCTTAACTGCAAAAGTTCACCTTTTGTATGTC 2766

QY 421 AACGCTATCGTTTGTCTACTGCTGGTGGTGGTGGTCTTTCCTTAACTCTGATAGTAC 480
   |||||
Db 2767 AATCTTTTCTCTTACGGTTCCTCTGCGCTCTCTGCTGCTCAACACTGATTCAGAA 2826

QY 481 AAGCTTGAAACAGACACACAAAGAAATATGTTGTTGGTTCATCATGACTCTTGGTGA 540
   |||||
Db 2827 AACAACAACAAATGTATCTAGAGTACAATATGTAATGTTGGTTCATCTGTACCACTT 2886

QY 541 GCTCTTCTCTATGGTTTTATATGTCACCTTGTGCGAGCTTTCTTCAAGAAATCTGTCAG 600
   |||||
Db 2887 TCGCTGGGATGGATTTGGTACTATCTCTGATACACTGCTCTTCAGGAAGTTTTCAG 2946

QY 601 CGAATCACTATACGCTCGCGCTCGAGTTCCAGATGGTCTTATGCTTTGCTGCCACTTGT 660
   |||||
Db 2947 AAGCATACATCTCAGCAGCTCTGGACTTGGCCAAATACCACTCTAGTTGCACTTGT 3006

QY 661 GTCTGCTCGTGGGATGTAGCGCTGCGGATTTCAAGGTGATAGCAGGAGAAAGCAAGA 720
   |||||
Db 3007 GTGTGACTCATAGGACTGTTTGAAGTGGAGAGTGGAGAACTCTGCCAAGTGGATGAGA 3066

QY 721 GATTTTATAGCTGGAGAGCTTTTGTACTATGTTGTTGTTGTTGTTGTTGTTGTTGTTG 780
   |||||
Db 3067 AACTACAAACTGGGAAAGTGTATATATCTTTGACTTTGGCTCTCAGCAGCTATTTCTGG 3126

QY 781 CAAGCATTTTGTGGGAGCTATTGGGTTGATCTTCTGTGCACTGCTCTCTGCTCTGGA 840
   |||||
```

```
RESULT 7
US-10-424-599-102519
; Sequence 102519, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 102519
; LENGTH: 2214
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_63592C.1
US-10-424-599-102519
```

```
Query Match      15.7%; Score 165.2; DB 16; Length 2214;
Best Local Similarity 51.8%; Pred. No. 1.9e-41;
Matches 374; Conservative 0; Mismatches 348; Indels 0; Gaps 0;

QY 245 TCTTATCGCGCTATCGTGTGTTTGTCTGCGGATTTGACAAATTAACCTCTACTCTT 304
   |||||
Db 623 TCTTTGATATATTTTGTCTTTGGAGTCTTAATGCTGTCACAAATATGATGACTCA 682

QY 305 ACGGTTAGCTTATATCCCTGTTTCTACTGCGTCTTTGATCATCTCCGCAATTAAGCT 364
   |||||
Db 683 CTGGACTCTTATACCTCTCGGCTTCTACCTATTGCTGATTTGTGTCATCACAGTTAGCTT 742

QY 365 TCACCTGCTCTTTTGGCAATTTTATGTTGAAGAAAGTTCACACCTTTCACTATAAAG 424
   |||||
Db 743 TTAATGCAAGTTTCTCATATTTTATCAATTTCTCAAAAGTTTCACTGCTCTGATTAATA 802

QY 425 CTATGCTCTCTTTCACGTGCTCGGCTGCTGCTTTCCTTAACTCTCATGATGACAGC 484
   |||||
Db 803 CTACAGTGGTCTTCTATCTGCTGCTGCTCTTCTGCTGCTTAAAGCAAGATGATGAAC 862

QY 485 TTGCAACGAGACACACAAAGGAATATGTTGTTGGTTCATCATGACTCTTGGTGCAGCTC 544
   |||||
Db 863 CATCTGTTTCTCAAGGAAAGTACATATTGTTGTTTCTATGATACCTCTGGAGCTTCTG 922

QY 545 TTCTTATGGTGTATATATGACCTTGTGAGCTTTTCTTACAAAGAAATCTGGTCAAGCA 604
   |||||
Db 923 CAGTGTACTCTTTTGTCTTTTCCCTCATGCACTGACCTTTGAGAAGGTTCTGAGAAGG 982

QY 605 TCAGTATACGCTCGGCTCGAGTTCAGATGCTTATGCTTTTGTGCTCCCACTTGTGCT 664
   |||||
Db 983 AATCATTTCTGTTGTTTGGAAATGCAAAATCTACACATCAATCTGTTGCTCTGTTGCTT 1042

QY 665 GCCTCGTGGGATGCTAGCCGCTGCGATTTTCAAGGTGATAGCAGGAGAAAGAGATT 724
   |||||
```

Db 1043 CTGTCATAGGCCCTATTTCGAAGTGGGAAATGGCGTACTTTTCATGAGAAATGAGGGTT 1102  
Qy 725 TTAAGCTTGGAGAGCTTTTGTACTATGTGGTGAATGTGTTTCAAGCCATAATCTGGCAAG 784  
Db 1103 TTCGAAAGGATATGTTTGCCTTATGTATGACTTTTGGTTTGGACTTCAATAGCCTGGCAGG 1162  
Qy 785 CATTTTTTGTGGAGCTATTGGGTTGTATCTTCTGTGCAATCGTCTGTGCTCTCTGGAATTA 844  
Db 1163 TATGCTCTGTGGTGTGGTGTGGCTGTGATCTTCTAGTGTCTTCTCTACTTCCAAATGTTA 1222  
Qy 845 TGGTCAGTGTCTGTCTCCGGTGAAGGTGATCTTTGGCCGTCAATTTGCTTTCAGGAGAAGT 904  
Db 1223 TAAGCACAGTTTCTTTAGCCGTAACCTCTATTGCTGTGTTATAGTTTTTTCATGATAAGA 1282  
Qy 905 TTCAGCCGGGAAAGGTGTGGTCTCTCCCTCTGGGGATCAGTCTCTTATTTCT 964  
Db 1283 TGAATGGGGTGAAGATAATTTCTATGCTTTTGGCTCTATGGGGTTTTGGCTCTTATATTT 1342  
Qy 965 AT 966  
Db 1343 AT 1344

## RESULT 8

US-10-437-963-82061/c  
; Sequence 82061, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 82061  
; LENGTH: 1513  
; TYPE: DNA  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_81525C.1  
US-10-437-963-82061

Query Match 14.4%; Score 151; DB 17; Length 1513;  
Best Local Similarity 48.0%; Pred. No. 5.3e-37;  
Matches 430; Conservative 0; Mismatches 465; Indels 0; Gaps 0;  
Qy 71 CTCATATGATCGTCTTACTTCCAAATGTGGGAAAGGATCTGTTTCCAAAGCTTCC 130  
Db 1200 CACTTCTTGGGAGTGTACTACAAATCAAGCGGCAATGAGGATGTCCACATTCG 1141  
Qy 131 TTCAAACGTTGGTTGTCCACTCAFTTTCTTCCCTCTTCTCTTATCTTTCTCCGCGTC 190  
Db 1140 TCCAAACTGTGGCTTCCGATTTTGTTCATTGCGCTATTTCTTTCCATTCAAAGACAT 1081  
Qy 191 GTCGTGTGCTTGAGAAACAAGACGACTCAATTTTCTCTCATGAAACCTCTCTCTTTA 250  
Db 1080 CTTCTACAAACACGACTAGTAGTCTCGCCCTTACAAATTTCTATCCCAAAATTAATC 1021  
Qy 251 TCGCGCTATCGTGTGTTTGTGCGGGATTTGACAATTTACTCTTACTCTTACGGGT 310  
Db 1020 TGATATATGTTGCTCGGCCCTCATCATGTGAGACGACTTGATGATTTCTATGGCC 961  
Qy 311 TAGCTTATATCCCTGTTTCTACTCGGTCTTTGATCATCTCCGCGCAATTAGGCTTCACTG 370  
Db 960 TACTATATCTTCCGGTCTCAACATATTCGCTCATCTGTGCTAGTCAGCTTGCCTTCAATG 901

Qy 371 CTCCTTTGTCATTTTATGGTGAAGCAAAAGTTTCACACCTTTTCACTATATAACGCTATCG 430  
Db 900 CTGCTTCTCATATATTTCTCAATGCTCAAAAATTCACCCCTCTGATTTTCAATTTCCGTAG 841  
Qy 431 TTTTGCTCACTGGTGGTGGCTAGTCTCTTCCCTTAACTCTGATAGTGACAAGCTTGCAA 490  
Db 840 TCCTCTTACGTTTCTGCTTCACTCTTGGAGTTGATGAAGATTTCTCAGGAACTACTA 781  
Qy 491 ACAGACACACAAGGAATATGTTTGGGTTTCATCATGACTCTTGGTGCAGCTTCTCTCT 550  
Db 780 GTATATCACATGGGAAGTACATTTTGGGTTTCTCTGTGACACTAGGGGCATCAGTACAT 721  
Qy 551 ATGGGTTTATATGCCACCTTGTGAGCTTTCTTACAGAATCTGGTCAGGAATCAGT 610  
Db 720 ACTGCTCATTTCTCTCCCTGATGCAAGTCACATTTGAGAAGGTTATTTAAGAGGGAGACCT 661  
Qy 611 ATAGCTCGCGCTCGAGTTCAGATGGTCTTATGCTTTGCTGCGCCTTGTGTCTGCCCTCG 670  
Db 660 TCTCAGTTGTGTGAACATGCAGATATATACAGCTCTCGTGGCAACATTGGCTTCTCTTG 601  
Qy 671 TGGGATGTAGCGCTGGCGATTTTCAAGGTGATAGCAGAGAGCAAGAGATTTTAAAGC 730  
Db 600 TTGGGTTATTTGCAAGTGGTGAATGGATGACTTTTACAAGGAGAGATGCATGCATTCAGT 541  
Qy 731 TTGGAGAGTCTTTGTACTATGTGGTGTGATCTTCTGTGCATCGTCTCTGGTCTCTGGAATATGTC 850  
Db 540 CTGGAAAGCTGTCTATGTAAATGACACTGCTGTGGACGGCTATATCTTGGCAGGTAGCAT 481  
Qy 791 TTGTGGAGCTATTTGGGTTGATCTTCTGTGCATCGTCTCTGGTCTCTGGAATATGTC 850  
Db 480 CAGTTGGAGTGGTGGGATTCATCTTGTGTGTGTCATCGCTGTTTCAAAATGTGATAAGCA 421  
Qy 851 GTGCTGTGCTTCCGGTGACCGTGATCTTGGCGCTCATTTGCTTCCAGGAGAAGTTTCAGG 910  
Db 420 CCCTAGCTCTACCCCATCATTCCTGTTTGTGCTGTGATTTTCTTTCATGACAAGATGGATG 361  
Qy 911 CGGGAAAGGTGTCGCTTTGCTCTCTCCCTCTGGGATCAGTCTCTTATTTCTA 965  
Db 360 GAGTAAAGATTATAGCTATGCTGATGGCCATTTGGGGATTTATGTATATGGCCA 306

## RESULT 9

US-10-425-114-13687  
; Sequence 13687, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E.  
; APPLICANT: Tabaska, Jack E.  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(5313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 13687  
; LENGTH: 1296  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB143-024-D8\_FLI  
US-10-425-114-13687

Query Match 13.8%; Score 144.4; DB 16; Length 1296;  
Best Local Similarity 52.4%; Pred. No. 6.3e-35;  
Matches 372; Conservative 0; Mismatches 326; Indels 12; Gaps 2;  
Qy 4 ATGAAGTGAAGACAGTTCTTGTAAATCAATAAAGTATATTTTGGCCATTTGGAATGTG 63

Db 70 AAGACGATGCGCGCTCTCTGCTGGCGACCAACTGCGGTGATGCTCGCGCTTGGCGTCAAC 129  
Qy 64 GGAGGCGCTCTAATGATCGCTCTACTTCCAAAATGGTGGCGAAGAGTCTGGTTTCCA 123  
Db 130 GGTGGGCGCTCTCAGCGCGCTCTACTTTAGCAAGGCGGACACCGCAGTGGCTCTCC 189  
Qy 124 AGCTTCCTTCAAACCGTTGGTGTCACTCATTTTCTTCCCTCTCTCTTATCTTTCCTC 183  
Db 190 GSAATGGCTCAGACCGGTGGCTGGCGCTCTGCTCCCGCAGTGGCGGCTTCTTACGTC 249  
Qy 184 CGCGCTGCTGTTGCTTGAAGAACAAAGAACACCTCCATTTTCTCTCATGAACCTCCT 243  
Db 250 CGCGCGAGCGCGC---GACCGAGCGCCCGGTGCTGCTACCCAGACCCAGCGCGC 306  
Qy 244 CTCTTATCGCCGCTATCTGTTGTTGCTCGTGGGATTTGACAAATACCTCTACTCT 303  
Db 307 ATACTGTCGCGCGCGGCTCGGCTCATCGCGGCGTGGCAACTTGTCTTAGGCC 366  
Qy 304 TACGGTTAGCTTATATCCCTGTTTCTACTGCGTCTTTGATCATCTCGGGCAATTAGGC 363  
Db 367 TGGGGCTCGAGTTCTGCGCGCTCTCCACCTCGCGATCCTCATCTCACGCACTGGCC 426  
Qy 364 TTCACCTGCTCTTTTGCATTTTATGGTGAACAAAGTTACACCTTTCACTATAAAC 423  
Db 427 TTACCGTGTCTTTCGGGTACCTGATCGTGGCGAGCGGTGACGATGGCGGTGAAC 486  
Qy 424 GCTATCGTTTGTCTACTGCTGCTGCTAGTCTTTGCGCTTAACTCTGATAGTGACAAG 483  
Db 487 GCGGTGGCGTGTGACGGTGGCGCGTCTGCTGGGCTGATGCTCTCTGACCGC 546  
Qy 484 CTTGCAACGAGACACAAAGGAATATGTTGTTGGTTCATCATGACTCTTGGTGAGCT 543  
Db 547 CCGCGCGGCTCACACGAGTCACTGCTGGGCTTCACTGCTCACCTGCGCGCGC 606  
Qy 544 CTCTCTATGGTTTATATGCACTTTCGAGCTTCTTACAAAGATC-----T 594  
Db 607 GTGCTGTACGGGTGTTCTGCGCTGGTGGAGCTCACCTACAGTGGCGCGGGGCG 666  
Qy 595 GGTGACGGAATACGCTATACGCTCGCGCTCGAGTTCAGATGGTCTTATGCTTGTGCG 654  
Db 667 GGAGGCGCGTGTGAGCTACGCGCTAGTGGTGGAGTTCGAGCTGTGATGGGTTCTGCGC 726  
Qy 655 ACTTGTGCTGCTCGTGGGATGCTAGCGGCTGGCGGATTTCAAGGTGAT 704  
Db 727 ACCGCGTTCTGACCGCTGGCATGATCGTCAACAGGACTTTCAGGTGAT 776

RESULT 10  
US-10-437-963-91425  
; Sequence 91425, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yinhua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 91425  
; LENGTH: 1177  
; TYPE: DNA  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_90000C.1  
US-10-437-963-91425

Query Match 13.5%; Score 142; DB 17; Length 1177;  
Best Local Similarity 48.4%; Pred. No. 3.5e-34;  
Matches 469; Conservative 0; Mismatches 485; Indels 16; Gaps 2;  
Qy 22 CTTGTAATCATAACTGTATATCTTGGCCATTTGGAAACTGTGGAGCCCTCTAAATGATG 81  
Db 142 CTGCTGGTCAATTTAGCGCGTGCCTCTCTCTCGCGCGCGCGCGCTCTCTCTCTC 201  
Qy 82 CGTCTCTACTTCCAAAATGGTGGCGAAGATCTGTTTCCAAAGCTTCTTCAAAACGGTT 141  
Db 202 CGCGTCTACTTCTGTGACCGCGGACCCGGCTGTGGCTGTCCGCCACGCTCCAGATCTCC 261  
Qy 142 GGTGTGCACTCATTTTCTTCTCTTATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 201  
Db 262 GGTGTGCGCTCTGCTGCGCGCTGTGCGTGTGCTCTACCGCGCGCGCAG----- 314  
Qy 202 GAAGAACAAAGAACTCACTTTTCTCATGAACCTCTCTCTTATTCGCGCGTATC 261  
Db 315 -----GCACGGATCGCAACCTCTCTCTCGCGCGCGCTCTCTCGCGCGCGCGC 366  
Qy 262 GTTGTGTTGCTCTGCGGATTTGACAAATTAACCTTACTCTTACGGGTTAGCTTATATC 321  
Db 367 GTGCTGGCGGCTGTACCGCGTGTGCTCTGTGTACCGCTGGGCTGGAGGCGCTG 426  
Qy 322 CCGTGTCTACTGCTCTTGTGATCATCTCGCGCAATTAGGCTTCACTGCTCTCTTTTGA 381  
Db 427 CGCGTGTCCACGCTGCTGCTGTGGCGACGAGCTGGCCTTCAACCGCGCTTTCGCG 486  
Qy 382 TTTTATGTTGTAAGAAAGTTTCACTTTTCACTATAAAGCTATCTGTTTGTCTCACT 441  
Db 487 TTTCTCTTCTGCGGCTCTCGGTTTCACTGCGGCTTCTCGGCAAGCGCTCTGCTGCTCACT 546  
Qy 442 GGTGTGCGGTAGTCTTCCCTTAACTCTGTATGTGACAACTTGCACAAACGACACAC 501  
Db 547 GTGAGGAGCGGTGTCGAGGAAGCGGTGTAAAGCGGTGGCGGACCCGATACCGGTGA 606  
Qy 502 AAGGAATATGTTGGGTTTCACTGATCTTGTGTGAGCTCTTCTTATGGGTTTATA 561  
Db 607 GGTGTGCACTTAGTGCAGCTCCAGCAAGTGGTCACTCTACAGTCCGAGATTTGGGATAG 666  
Qy 562 TTGCACTTGTGCGAGCTTTTCAAGAAATCTGTCAGCGAATCACTATACGTCGCG 621  
Db 667 ATGAGTGTGGCGGCTTCTAGGTGGAATTTGCGGCTTACGAGCCCTTACGCGAGGTG 726  
Qy 622 CTCGAGTTCAGATGCTTATGCTTCTGCTGCACTTGT-GTCTGCTCTGCGGATGCT 680  
Db 727 ATGAGATGAGCGCGGTGATGGCGCGGACACCGCGGTGTGCTGCTCGGATGGC 786  
Qy 681 AGCGCTGCGGATTTCAAGGTGATAGCAGGAGAAAGAGATTTTAAAGCTTTGGAGAGTC 740  
Db 787 GATCAAGGCGCTTTCCAGCGGTGCGCGGGAAGCGGCGGTTCGGGCTCGCGCGCGC 846  
Qy 741 TTTGTACTATGTTGATTTGTTTCAAGCAATCTGTCAGCGAATCTGTCGCAAGCATTTT 800  
Db 847 CAACTACTACCTGCTCTGCGCTGGGACCGCTGTGTCGTCGAGCTGCTCAACCTGGCAT 906  
Qy 801 TATTGGGTTGATCTTCTGTGATCTGCTCTGCTCTGGAATATGTCAGTCTGCTCTGCT 860  
Db 907 CATGGGCTCATCACTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 966  
Qy 861 TCCGCTGAGCGGTGATTTGGCGCTCATTTGCTTCCAGGAGAGTTCAGCGCGGGAAGG 920  
Db 967 GCGGCTCTCGCAGGCTCTCGCGCTCATCTTCTTCCAGGAGTTCGACGGGACGAGGG 1026  
Qy 921 TGTGCTTGGCTCTCTCCCTCTGGGATCAGTCTCTTATTTCTATGACAGGTTAAATC 980  
Db 1027 CATCGGCTCGGCTCTGCTCTGGGATTCGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1086  
Qy 981 CGAGGAGAG 990  
Db 1087 GAAGAAGAG 1096

## RESULT 11

US-10-416-898-9

; Sequence 9, Application US/10416898

; Publication No. US20040172670A1

; GENERAL INFORMATION:

; APPLICANT: Walker, Elabeth

; APPLICANT: Dellaporta, Stephen

; TITLE OF INVENTION: MAIZE YELLOW STRIPEL AND RELATED GENES

; FILE REFERENCE: 44574-5108-US

; CURRENT APPLICATION NUMBER: US/10/416,898

; CURRENT FILING DATE: 2003-05-16

; PRIOR APPLICATION NUMBER: PCT/US01/43101

; PRIOR FILING DATE: 2001-11-16

; PRIOR APPLICATION NUMBER: US 60/249,222

; PRIOR FILING DATE: 2000-11-16

; NUMBER OF SEQ ID NOS: 20

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 9

; LENGTH: 83698

; TYPE: DNA

; ORGANISM: Zea mays

US-10-416-898-9

Query Match 12.7%; Score 133; DB 17; Length 83698;

Best Local Similarity 48.0%; Pred. No. 3e-30;

Matches 455; Conservative 0; Mismatches 475; Indels 17; Gaps 2;

```
Qy 20 TTCCTGTAATCATAACTGTATATCTTGGCCATTTGGAACATGTGGAGGCCCTCTAATGA 79
   |||||
Db 67235 TTCTGTTTTCATAGCACTCTCTCTCATCTCTGCTCAAGCCATTTCTGTTCTTTG 67294

Qy 80 TCGCTCTCTACTCCAAATGTGCGAAGATCGTGTTCCTCAAGCTTCTTCAACCG 139
   |||||
Db 67295 GTCCGTTTATTACAAATGAAGTGAACAGTAATGATCTACTCTTGTTCAAACTG 67354

Qy 140 TTGG-TTGTCCACTCATTTCTTCCCTCTCTCTTATCTTCTCCCGCGTCTGCTGTC 198
   |||||
Db 67355 GTGGCTTTCCCATCTTATCTCCCTCTTCTCTCTCTCTCTCTCTCAATCTTCTT 67414

Qy 199 CTTGAAGAACAGAAACGACTCCATTTTTCCTCATGAACCTCTCTCTTATCGCCGCT 258
   |||||
Db 67415 CTT-----CTTCTTCTTCTTCTTCTTCTTCAAGACTCTGGTTGGATTTAT 67458

Qy 259 ATCGTTGTTGGTTTCTCGTGGATTTCACAAATACCTCTACTCTTACGGGTAGCTTAT 318
   |||||
Db 67459 CTTTCTCTGGTTTGTCTATTGTTTATGATTAATTTTATATCTCTGTTGGACTTTTGTAT 67518

Qy 319 ATCCCTGTTTCTACTCGCTCTTTGATCATCTCCGCGCAATTAGGCTTCACTGCTCTCTT 378
   |||||
Db 67519 CTCTCTGCTTCAACTTATTCGATTTCTATGTGCTTTCACAGTTAGCTTTCATGGTGTCTTC 67578

Qy 379 GCATTTTATTTGGTGAAGAAAGTTACACTTTTCACTATAAAGCTATCGTTTGTCTC 438
   |||||
Db 67579 TATTATTACATCAATTTCTCAGAAAATCACTTGTGTGATTTTCTTCTCAGTGTGTCTC 67638

Qy 439 ACTGTTGTTGGTGTAGTCTTGGCTTAACTCTGATAGTACAGCTTGCAGACAGACACA 498
   |||||
Db 67639 TCTATCTCTGCTGTGTGGTTTCTTGACGATGATTCAAATAGCCCATCAGGAGATTCT 67698

Qy 499 CACAAGGAATATGTTGTTGGGTTTCATCATGACTCTTGTGTGAGCTTCTTCTATGGGTTT 558
   |||||
Db 67699 AAGTGGAGTACTTGATTTGGGTTTCTGTCAGTTTGTCTTCTTATCTATCTTCTT 67758

Qy 559 ATATTGCCACTTGTGCGATTTCTTAAAGAAATCTGTTGTCAGCAATCACGTATACGCTC 618
   |||||
Db 67759 CAGCTCTCTCTATGTCAGTTTCTTTCGAGAAGGTTCTCAAGAGTGAAGTCTCTCTATG 67818

Qy 619 GCGCTCGAGTTCCAGATGTTCTTATGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 678
   |||||
Db 67819 GTTCTCGAGATGCMAATCATACGTCGCTTGTGGCTTCTTGTGTAGCGGTTATCGGATTG 67878
```

## RESULT 12

US-10-437-963-72452

; Sequence 72452, Application US/10437963

; Publication No. US2004012343A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; APPLICANT: Wu, Wei

; APPLICANT: Boukharov, Andrey A.

; APPLICANT: Barbazuk, Brad

; APPLICANT: Li, Ping

; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With

; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53221)B

; CURRENT APPLICATION NUMBER: US/10/437,963

; CURRENT FILING DATE: 2003-05-14

; NUMBER OF SEQ ID NOS: 204966

; SEQ ID NO 72452

; LENGTH: 789

; TYPE: DNA

; ORGANISM: Oryza sativa

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_72830C.1

US-10-437-963-72452

```
Query Match 12.4%; Score 130.4; DB 17; Length 789;
Best Local Similarity 53.1%; Pred. No. 1.4e-30;
Matches 306; Conservative 0; Mismatches 261; Indels 9; Gaps 1;
```

```
Qy 20 TTCTTGTAAATCAATAACTGTATATCTTGGCCATTTGGAACATGTGGAGGCCCTCTAATGA 79
   |||||
Db 218 TCCTCGTGGGCTCAACTGCGGATGCTGTTGGTGGCCACCGGCGGCGCGATCCTTA 277

Qy 80 TGGCTCTCTACTTCCAAAATGTGGGAAAGATCTGTTTCCAAAGTCTCTTCAAAACCG 139
   |||||
Db 278 GCCGGCTCTACTTTCAGCAGGCGGCCACAGGAAGTGGCTCTCGGGTGGCTCGAGACCG 337

Qy 140 TTGGTTGTCCACTCAATTTTCTTCCCTCTTCTTATCTTTCCTCCGCGCTCGTCTGTTGCC 199
   |||||
Db 338 CCGGGTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 397

Qy 200 TTGAGAACAGAAACGACTCCATTTTCTCTCATGAACCTCTCTCTCTTATCGCCGCTA 259
   |||||
Db 398 ACCGCGCG-----CGCCGCTGTTCTCTACGCGCGCGCGCTGCTGCTGCGGCGC 448

Qy 260 TCGTTGTTGTTGCTCGTGGGATTTTGCATAATACCTCTACTCTTACGGGTTAGCTTATA 319
   |||||
Db 449 CGGTGCTCGGGTCTTCCACGGCGCCGACGACTCTGCTACGCTACGCGCTCGCCTACC 508

Qy 320 TCCTGTTTCTACTGCGTCTTTGATCAITCTCCGCGCAATTAGGCTTCACTGCTCTCTTTG 379
```

Db 509 TGGCGGTGCCACCTCCGCCATCTCTCATCCACGAGCTCGCGTTACCGTCTTCTTCG 568  
Qy 380 CATTTTTTATGTTGAAGCAAAAGTTTCAACACCTTTTCACTATAAAGCGTATCGTTTGTCTCA 439  
Db 569 CGTGCCCTCATCGTGGCGAGCGGCTACCGCGCGAGCGTGAACGCCGTGGCGTGTCTCA 628  
Qy 440 CTGGTGGTGGCGGTAGTCCCTTGGCCCTTAACTCTGATAGTGAAGCTTGGCAACGAGACAC 499  
Db 629 CTATCGGCGCGGTGCTGGGCTCCACGCGTCCAAAGGACCGACCCGCGGGGTGACCA 688  
Qy 500 ACAAGGAATATGTTGTTGGGTTCAATCATGACTCTTGGTGCAGCTCTTCTCTATGGTTTA 559  
Db 689 CCGGAAGTACTGGATGGGCTTCTCTCACTCCGCGCGCGCGCTGTATCGGCGTCA 748  
Qy 560 TATTGCCACTTGTCCGAGCTTTCCTTACAAGAAATCTG 595  
Db 749 TCTTGGCGTCTGCGAGCTCGGTACAGACGCCG 784

## RESULT 13

US-10-424-599-31364  
; Sequence 31364, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 31364  
; LENGTH: 1373  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_128327C.1  
US-10-424-599-31364

Query Match 12.4%; Score 130.2; DB 16; Length 1373;  
Best Local Similarity 48.5%; Pred. No. 2.3e-30;  
Matches 360; Conservative 0; Mismatches 383; Indels 0; Gaps 0;

Qy 258 TATCGTTGTTGTTGCTCGTGGGATTTGACAAATTACCTCTACTCTTACGGGTTAGCTTA 317  
Db 500 TATTGTTTGGGTTTCTTAAGTGTCTGTGATAACCTCATGTATGCTTATGCTATGCTTA 559  
Qy 318 TATCCCTGTTTCTACTGCGTCTTTGATCATCTCCGCGCAATTAGGCTTCACTGCTCTTT 377  
Db 560 CTCTCCCTGCATCCACCGCTCTACGTTGCTTCAATCCCTTGTGTTTCTGCACTCTT 619  
Qy 378 TGCATTTTTTATGTTGAAGCAAAAGTTTCAACCTTTTCACTATAAAGCGTATCGTTTGT 437  
Db 620 TGGATACTTTTCTGTGAAAAACAAGTGAATGCTTCGATAGTAATTCGTTTTCATCAT 679  
Qy 438 CACTGGTGGCGGTAGTCTTGGCCCTTAACTCTGATAGTGAAGCTTGGCAACGAGAC 497  
Db 680 AACCAATTGCAATGACCATCATTTGCACTGGACTCGAGTTTCAGACAGATATGCCAATCAG 739  
Qy 498 ACACAAGGAATATGTTGTTGGGTTTCATCATGACTCTTGGTGCAGCTCTTCTATGGGTT 557  
Db 740 TGACAGTAGTACATCATGGGATTTGTGFGGATGTTTATGATCTGCTTTCACGGGCT 799  
Qy 558 TATATTGCCACTTGTGAGCTTTCTTACAAGAAATCTGGTCAGCGAATCAGCTATACGCT 617  
Db 800 TATATTGCTCTCTCGGAGCTTGTCTTGTGAAGTTGCTTGAAGAGATCTTCTCATCGT 859  
Qy 618 CCGGCTCGAGTTCCAGATGGTCTTATGCTTTCGCTGCGCACTTGTGCTCGCTCGTGGGAT 677

Db 860 TGTCTGAGCAGCAAGTCATGGTTTCTCTGTTTCCATTTCTGTTTACCACTGTAGGGAT 919  
Qy 678 GCTAGCCGCTGCGGATTTTCAAGGTGATAGCAGGAGCAAGCAAGATTTTAAAGCTTGGAGA 737  
Db 920 GATTATGATGTTGATTTTCCAAAGGATGGCATGAGGCTACCCTTTCAAAGGTGGTAG 979  
Qy 738 GTCTTTGTACTATGTGGTGAATTTTCAACGCGCAATAATCTGGCAAGCAATTTTGTGGG 797  
Db 980 AAGTCTTATTATCTGTTATCATTTGGGGTGCATCACTTTTCAGCTGGGGGTTCTGGG 1039  
Qy 798 AGCTATTGGGTTGATCTTCTGTGCATCGTCTCTGTCTCTGGAATATGTCAGTCTCT 857  
Db 1040 GGGCACTGCTGTAATTTTCTTGGGCTCTACTGTGTAGCAGGTGCTTAATGCAAGTAA 1099  
Qy 858 GCTTCCGCTGACGGTGAATCTTGGCGCTCATTTGCTTCCAGGAGAAAGTTTACGCGCGGAA 917  
Db 1100 AACCCCATACAGCATTTGACGTGTTTACTGCTAAAGGACCCTATGAGTGGTTTCA 1159  
Qy 918 AGGTGCTGCTTGGCTCTCTCCCTCTGGGATCACTCTTATTTCTATGGACAGTTAA 977  
Db 1160 GATCCTCTCCCTAGTGATCACCTTTTGGGATTTGGCTCATATATTTATGGAAGTTCTAA 1219  
Qy 978 ATCCGAGGAGAACTAAGGCTC 1000  
Db 1220 AATATGGCCATAGACTCCAGTTC 1242

## RESULT 14

US-10-425-114-10574  
; Sequence 10574, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 10574  
; LENGTH: 1333  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: 700944981\_FLI  
US-10-425-114-10574

Query Match 12.1%; Score 127; DB 16; Length 1333;  
Best Local Similarity 48.7%; Pred. No. 2.4e-29;  
Matches 346; Conservative 0; Mismatches 365; Indels 0; Gaps 0;

Qy 258 TATCGTCTGTTGTTGCTCGTGGGATTTGACAAATTACCTCTACTCTTACGGGTTAGCTTA 317  
Db 462 TATCGTTTGGGTTTCTTAAGTGTGCTGATAACCTCATGTATGCTTATGCCATGCTTA 521  
Qy 318 TATCCCTGTTTCTACTGCGTCTTTGATCATCTCCGCGCAATTTAGGCTTCACTGCTCTCT 377  
Db 522 CTCTCCCTGCATCCACTGCTTCACTGTGGCATCATCATCCCTTGTGTTTTCGGCGCTCTT 581  
Qy 378 TGCATTTTTTATGTTGAAGCAAAAGTTTCAACCTTTTCACTATAAAGCTATCGTTTGTCT 437  
Db 582 TGGATACTTCTTGTGAAGAAACAAAGTGAATGCTTCAATAGTGAATTCGCTTTTTCGTCT 641  
Qy 438 CACTGGTGGTGGCGGTAGTCTTGGCCCTTAACTCTGATAGTGCAGAGCTTGCAGGAGAC 497  
Db 642 AACCGCTGCATTTGACCATCATTTGCCCTGGACTCGAGTTTCAGACAGATATCCAGCATTAG 701  
Qy 498 ACACAAGGAATATGTTGTTGGGTTTCATCATGACTCTTGGTGCAGCTCTTCTATGGGTT 557

```
Db 702 TGACAGTGAATACATCATGGGATTTGATGGATGTTTATAGATCTGCTTCCATGGGCT 761
Qy 558 TATATTGCCACTTGTGAGCTTTCTTACAAGAAATCTGGTCAGCGAATCAAGTATACGCT 617
Db 762 TATTTTCGCTCTCTCAGAGCTCGTCTTTGTGAAGTTGCTCGGAAGAAGATCCTTTATCGT 821
Qy 618 CGCGCTCGAGTTCAGAGTGGCTTATGCTTTGCTGCCACTTGTGCTGCCCTGCTGGGAT 677
Db 822 TGTTCTGAGCAGCAGTCATGGTTTCTATTTGCAATTCGTATTACCATCTGATGGGAT 881
Qy 678 GCTAGCCGCTGGCGATTTCAAGGTGATAGCAGGAGAAGCAAGATTTTAAGCTTGGAGA 737
Db 882 GATTGTGAGTGTGATTTTCAAGGATGACATGAGGCTACCATTTCGAAGTGGTAG 941
Qy 738 GTCTTTGACTATGTGGTATGTTGTCACGGCCAAATCTGGCAAGCAATTTTGTGGG 797
Db 942 AAGTCTTATTATCTTGTATCATTTGGGGTGCATCACTTTTCAGCTGGGGTTCTGGG 1001
Qy 798 AGCTATTGGGTTGATCTTCTGTGCATCTCTCTGCTCTCTGCTCTGCAATATATGCTCT 857
Db 1002 GGGCACTGCTATAATTTCTTGGGCTCTACTGTGCTAGCAGTGTGCTTAATGCAATAG 1061
Qy 858 GCTTCCGGTGCAGTGTCTTGGCGTCAATTTGCTTCCAGGAGAAGTTTTCAGGGGGGAA 917
Db 1062 AACACCCATAACAAGCATTTGCAGCTGTTATCTGCTAAAGGACCCCTATGAGTGGTTCAA 1121
Qy 918 AGGTGTGCTGTTGGCTCTCTCCCTCTGGGATCAGTCTCTTATTCTATGG 968
Db 1122 GATCCTCTCCCTAGTAGTACACCTTTTGGGGATTTGGCTCATATATTTATGG 1172
```

## RESULT 15

US-10-437-963-43860/c

```
; Sequence 43860, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Brad
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 43860
; LENGTH: 2381
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(2381)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_46977C.1
US-10-437-963-43860
```

```
Query Match 11.3%; Score 118.8; DB 17; Length 2381;
Best Local Similarity 47.6%; Pred. No. 1.4e-26;
Matches 351; Conservative 0; Mismatches 387; Indels 0; Gaps 0;

Qy 259 ATCGTTGTGGTGTGCTGCGGGATTTGACAAATTAACCTTACTCTTACGGGTTAGCTTAT 318
Db 2021 ATCGTCTGGGCTCATCATCGCCGCGACGACATGATGATGCTGTGGCTCAAGTAC 1962
Qy 319 ATCCCTGTTTACTCGGTCTTTGATCATCTCCGCGAATTAAGCTTTCACTGCTCTCTT 378
```

```
Db 1961 CTCCCCGCTTCGACCTACTCGCTCATCTGTGCCAGCAGCTCGCGTTCAATGTGCTCTTC 1902
Qy 379 GCATTTTATGTTGGAAGCAAAAGTTTCAACCTTTTCACTATAAAACGCTATCGTTTGTCTC 438
Db 1901 TCATACGTGTCTCAACTCCAGAAAGTCACTCTCTGATATTTCAAATCCGTCTGCTGCTC 1842
Qy 439 ACTGGTGGTGGCGTGTGCTTCCCTTAACTCTCATAGTAGTACAAAGCTTGGAAAACGAGACA 498
Db 1841 ACCATGTCGCTTCGCTCATCGGAGTCAGCAAGAGTCTCAGGGGGTACACGGCGTCTCG 1782
Qy 499 CACAAGGAATATGTTGTTGGGTTTCATCATGACTCTTGGTGCAGCTCTTCTCTATGGGTTT 558
Db 1781 GGAGGGAAGTATCTGCTCGGTTTCGCTGACGCTGGGGGGTCTGTCACCTACTCTCGCTG 1722
Qy 559 ATATTGCCACTTGTGAGCTTTCTTAACAAGAAATCTGGTCAGCGAATCAGATATACGCTC 618
Db 1721 ATCTTCGCGTGTATGTCAGCTCACCTTCGAGACCATCATCAAGAAAGCACACCTTCTCAGCC 1662
Qy 619 GCGCTCGAGTTCACAGATGGTCTTATGCTTTGCTGCCACTTGTCTGCTCGTGGGGATG 678
Db 1661 GTCTCAACATGTCAGATCTACACGGCCTCGTGGCGACGGCCGTCGTTGGTGGGCTG 1602
Qy 679 CTAGCCGCTGGCGATTTCAAAGGTGATAGCAGGAGAAGCAAGAGATTTTAAAGCTTGGAGAG 738
Db 1601 TTGCGAGCGCGAGTGGAGTCTGCTGAGGGGGGAGATGAACGCTTCAGTCTCGGGCAG 1542
Qy 739 TCTTTGTACTATGTGTGATTTGTTCACGGCCATTAATCTGGCAAGCAATTTTGTGGGA 798
Db 1541 TTCTCTACCTGATGACGCTGCTGTGGGGCGGCGTGTGTCGAGAGTGGCCAAACATCGGG 1482
Qy 799 GCTATTGGGTTGATCTTCTGTGCATCGTCTCTGTCTCTGAAATATATGTCAGTCTCTG 858
Db 1481 GTGCTCGGCTCATCTTCGAGGTGTCGGCGCTCTTCTCCAACGTGATCAGCACGGTCTCA 1422
Qy 859 CTTCCGGTGCAGGTGATCTTTGGCGCTCATTTGCTTCCAGGAGAAGTTTTCAGGGGGGAAA 918
Db 1421 CTGCGGTCTATCCGTTCTTTCGCGGTGTTGTTCCACGACAGGATGAACGGGTGAAG 1362
Qy 919 GGTGTGCTTTGGTCTCTCCCTCTGGGGATCAGTCTCTTATTCTTATGAGACAGTTTAAA 978
Db 1361 ATCGTGGCCATGCTGATTTGCAATTTGGGGATTTTATTTGCTATCTGTTTTCAGCACTATCTA 1302
Qy 979 TCCGAGGAGAAGACTAAG 996
Db 1301 GATGGCAAGAAAGCGAAG 1284
```

Search completed: November 1, 2004, 23:39:51  
Job time : 553.317 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 1, 2004, 05:45:29 ; Search time 3737.17 Seconds  
(without alignments)  
10228.408 Million cell updates/sec

Title: US-09-913-767-2  
Perfect score: 1049  
Sequence: 1 aaagatgaatgaagacagt.....actgattatgtagcttaaaa 1049

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_hic:\*  
4: gb\_est3:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_gse1:\*  
9: gb\_gse2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	999.8	95.3	1230	CNS0A99D	BX821376 Arabidops
2	521	49.7	827	BZ515292	BZ515292 BOMOL207R
3	515	49.1	527	AQ958362	AQ958362 LERAX127R
4	417.2	39.8	745	AQ958361	AQ958361 LERAX127F
5	414.4	39.5	1059	BX841065	BX841065 BX841065
6	410.8	39.2	688	BH589732	BH589732 BGS0504TR
7	334.4	31.9	463	BX836413	BX836413 BX836413
8	308.2	29.4	985	CK257466	CK257466 EST741103
9	306.2	29.2	986	CK243751	CK243751 EST727388
10	304.6	29.0	680	BG439880	BG439880 GA_Ba000
11	302.2	28.8	998	CK247918	CK247918 EST731555
12	301.4	28.7	1013	CK251896	CK251896 EST735533
13	297.2	28.3	907	CK257887	CK257887 EST741524
14	290.6	27.7	979	CK252005	CK252005 EST735642
15	288.4	27.5	988	CK248243	CK248243 EST731890
16	285.2	27.2	971	CK272858	CK272858 EST718936
17	281.8	26.9	645	AU238506	AU238506 AU238506
18	280.4	26.7	930	CK253429	CK253429 EST737066
19	278.8	26.6	929	CK249872	CK249872 EST73509
20	277.4	26.4	948	CK249994	CK249994 EST733631
21	274.6	26.2	908	CK247701	CK247701 EST731338
22	272.4	26.0	932	CK256828	CK256828 EST740465
23	268	25.5	948	CK257752	CK257752 EST741389
24	266.4	25.4	899	CK256093	CK256093 EST739730

25	266	25.4	928	7	CK245031	CK245031
26	265.2	25.3	688	1	AJ807511	AJ807511
27	262.6	25.0	894	7	CK247243	CK247243
28	261.8	25.0	868	7	CK245030	CK245030
29	258.6	24.7	882	7	CK258020	CK258020
30	258.6	24.7	889	7	CK251126	CK251126
31	256.8	24.5	880	7	CK257437	CK257437
32	255	24.3	892	7	CK246384	CK246384
33	250.8	23.9	784	8	BH601144	BH601144
34	249.6	23.8	722	4	BM436674	BM436674
35	242.8	23.1	359	8	BH441465	BH441465
36	239	22.8	832	7	CK259060	CK259060
37	236	22.5	703	6	CB006438	CB006438
38	233	22.2	829	7	CK262517	CK262517
39	226.4	21.6	827	7	CK249190	CK249190
40	222	21.2	902	7	CK262518	CK262518
41	220	21.0	665	6	CD714562	CD714562
42	219.8	21.0	588	5	BU762991	BU762991
43	215.2	20.5	616	5	BQ873555	BQ873555
44	215.2	20.5	701	5	BQ874152	BQ874152
45	215	20.5	668	6	CB008408	CB008408

ALIGNMENTS

RESULT 1  
CNS0A99D  
LOCUS  
DEFINITION  
Arabidopsis thaliana Full-length cDNA Complete sequence from clone GSLSIL372D07 of Silique of strain col-0 of Arabidopsis thaliana (thale cress).  
ACCESSION  
BX821376  
VERSION  
HTC; GSLT\_CDNA.  
KEYWORDS  
Arabidopsis thaliana (thale cress)  
SOURCE  
Arabidopsis thaliana  
ORGANISM  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
REFERENCE  
1 (bases 1 to 1230)  
Castelli, V., Aury, J.M., Jaillon, O., Wincker, P., Clepet, C., Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V., Temple, G., Caboche, M., Weissenbach, J., and Salanoubat, M.  
Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences: A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation  
JOURNAL  
Unpublished  
REFERENCE  
2 (bases 1 to 1230)  
Genoscope.  
TITLE  
Direct Submission  
JOURNAL  
Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)  
COMMENT  
The sequences are based on single pass reads.  
Life Technologies (a division of Invitrogen) members carried out full-length libraries construction : Temple G.  
Genoscope members carried out sequencing and annotation : Castelli V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M.  
URG INRA : Clepet C., Caboche M.  
Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich Information center for Protein Sequences). 5 prime and 3 prime are assembled with Phrap.  
http://www.genoscope.cns.fr/externe/sequences/Banque\_Projet\_EF/Full-length  
http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.

FEATURES  
source  
1. 1230  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/strain="Col-0"  
/db\_xref="taxon:3702"



301	Qy	TCTTACGGGTTAGCTTATATACCTGTTTCTACTACGGGCTTTTGATCATCTCCGCGCAATTA	360
504	Db	GCTTACGGGTTAGCTTACCTCCAGTTTCCACTTCTTTGGTCATCTCTCCCAACTG	445
361	Qy	GGCTTCACGTCTCTTTTGCAATTTTATGGTGAAGCAAAAGTTACACCTTTCACATA	420
444	Db	GCTTTCACGTCTTCTTGCCCTTTTTCATGGTGAAGCAAAAGTTACGCGCATTTATTATA	385
421	Qy	AACGCTATCTGTTTTCCTCACTGGTGGCGGTAGTCTTGTGCCCTTAACTCTCATAGTAGAC	480
384	Db	AACGCCGTGTTTTGCTTACTCTGGGTGGCGGAATCTTTGCCCTTCACACCGACGGTAGAC	325
481	Qy	AAGCTTGCAAACGAGACACACAAGGAATATGTTGTGGGTTTCATCATGACTCTTGGTGCA	540
324	Db	AAGCTTGCNAAAGACACACAAGGAGTATATCGTTGGGTTCTCATGACCTTGTCTGGG	265
541	Qy	GCTCTTCTCTATGGGTTTATATTTGCCACTTTGTCGAGCTTTCCTTACAAGAAATCTGGTCAAG	600
264	Db	GCTATTCTCTACCGCTTTGTTATTTGCCGCTTGTGGAGCTCACTTACAAGAAAGCTCGTCAA	205
601	Qy	CGAATCAGTATACGCTCGCGCTCGAGTTCAGATGGTCTTATGCTTTCCTGCCACTTGT	660
204	Db	CGAATCAGCTACACGCTTGTGCTCAGATGCAATTTGGTCTTGTGCTTTTGTGCCCACTTGC	145
661	Qy	GTCTGCCCTCGTGGGATGCTAGCCGCTGCGCAATTTCAAGGTGATAGCAGGAGAACGAAG	720
144	Db	TTCTGCCCTCGTTGGGATGCTAGCTGATGCGAATTTCAAGGAGTT-----AGCAAGA	94
721	Qy	GATTTTAAAGCTTGGAGAGTCT-----TTGTACTATGTGGTGAATGTGTTCACGGCCATA	774
93	Db	GATTTTAAAGCTTGGAGGCTTACTACGATATTACTATGTGTGGTGTGTTCACGGCCATC	34
775	Qy	ATCTGGCAAGCATTTTTTGTGGGAGCTATTGGG	807
33	Db	ATCTGGCAAGGCTTTTTCTTGGGATCTATTGGG	1

RESULT 3	LOCUS	527 bp	DNA	linear	GSS 28-JAN-2000
AQ958362/c	LOCUS				
	DEFINITION	AQ958362 LERAX12TR LERA Arabidopsis thaliana genomic clone LERAX12, genomic survey sequence.			

REFERENCE	1 (bases 1 to 527)
AUTHORS	Buell, C.R., Lin, X., Pai, G., Barnstead, M., Bowman, C., Utterbach, T., Feldblyum, T., Liang, F., Creasy, T. and Fraser, C.M.
TITLE	Genomic survey sequencing of <i>landisberg erecta</i> ecotype of <i>Arabidopsis thaliana</i> and identification of sequence-based polymorphisms
JOURNAL	Unpublished (2000)
COMMENT	Contact: Xiaoying Lin The Institute for Genomic Research 9712 Medical Center Dr., Rockville, MD 20850, USA Tel: 301 838 0200 Fax: 301 838 0208 Email: at@tigr.org For additional information, see <a href="http://www.tigr.org/tdb/at/at.html">http://www.tigr.org/tdb/at/at.html</a> Seq primer: TR Class: shotgun.

## FEATURES

[illegible]JOURNAL  
COMMENT

The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: [at@tigr.org](mailto:at@tigr.org)

For additional information, see <http://www.tigr.org/cdb/at/at.html>

Seq primer: TF  
Class: shotgun.

FEATURES  
source  
1..745  
Location/Qualifiers  
/organism="Arabidopsis thaliana"  
/mol\_type="genomic DNA"  
/ecotype="landsberg erecta"  
/db\_xref="taxon:3702"  
/clone="LERAX12"  
/clone\_lib="LERA"  
/note="Organ: Leaf; Vector: pHOS1; Total genomic DNA was sheared to 0.9-1 Kbp before ligation."

## ORIGIN

Query Match 39.8%; Score 417.2; DB 8; Length 745;  
Best Local Similarity 99.1%; Pred. No. 1.7e-113;  
Matches 419; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
Qy 1 AAGATGAAGATCAAGACAGTCTTGTATCATAACTGTATATCTTGGCCATTGGAAC 60  
Db |||||  
323 AAGATGAAGATGAAGACAGTCTTGTATCATAACTGTATATCTTGGCCATTGGAAC 382  
Qy 61 TGTGAGGCCCTCTAATGATCGCTCTCTACTTCCAAATGTTGGCGAAAGGATCTGGTT 120  
Db |||||  
383 TGTGAGGCCCTCTAATGATCGCTCTCTACTTCCAAATGTTGGCGAAAGGATCTGGTT 442  
Qy 121 CCAAGCTCTCTCAAAACGGTGTGTCACCTCAATTTCTCCCTCTCTTATCTTTC 180  
Db |||||  
443 CCAAGCTCTCTCAAAACGGTGTGTCACCTCAATTTCTCCCTCTCTTATCTTTC 502  
Qy 181 CTCGCCGCTCGTGTGCTTGAAGACAAGAACGACTCAATTTCTCTCATGAACCT 240  
Db |||||  
503 CTCGCCGCTCGTGTGCTTGAAGACAAGAACGACTCAATTTCTCTCATGAACCT 562  
Qy 241 CCTCTCTTATCGCGCTATCGTGTGTTGCTGCGGATTTGACAACTTACTCTAC 300  
Db |||||  
563 CCTCTCTTATCGCGCTATCGTGTGTTGCTGCGGATTTGACAACTTACTCTAC 622  
Qy 301 TCTTACGGGTTAGCTTATATCCCTGTCTACTCGCTTTTGATCATCTCGCGCAATTA 360  
Db |||||  
623 TCTTACGGGTTAGCTTATATCCCTGTCTACTCGCTTTTGATCATCTCGCGCAATTA 682  
Qy 361 GCCTTCACTGCTCTTTGCAATTTTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 420  
Db |||||  
683 GCCTTCACTGCTCTTTGCAATTTTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 742  
Qy 421 AAC 423  
Db |||||  
743 AAC 745

RESULT 5  
BX841065  
LOCUS  
DEFINITION  
BX841065 Arabidopsis thaliana Flowers and buds Col-0 Arabidopsis  
thaliana cDNA clone G51F65ZD02 SPRIM, mRNA sequence.  
ACCESSION  
BX841065  
VERSION  
BX841065.1 GI:42535148  
KEYWORDS  
EST.  
SOURCE  
Arabidopsis thaliana (thale cress)  
ORGANISM  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
1 (bases 1 to 1059)  
REFERENCE  
Castelli V., Aury J.M., Jallion O., Wincker P., Clepet C.,  
Menard M., Cruaud C., Quetier P., Scarpelli C., Schachter V.,  
Temple G., Caboche M., Weissenbach J. and Salanoubat M.  
TITLE  
Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences:  
A Combined Approach to Evaluate and Improve Arabidopsis Genome  
Annotation  
JOURNAL  
Unpublished (2004)

## COMMENT

Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: [seqref@genoscope.cns.fr](mailto:seqref@genoscope.cns.fr), Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr)  
The sequences are based on single pass reads.  
Life Technologies (a division of Invitrogen) members carried out  
full-length libraries construction : Temple G.  
Genoscope members carried out sequencing and annotation : Castelli  
V., Aury J.M., Jallion O., Wincker P., Menard M., Cruaud C.,  
Schachter V., Weissenbach J., Salanoubat M.  
URGV INRA : Clepet C., Caboche M.  
Annotation is based on the June 2003 version of the Arabidopsis  
genome released by MIPS (Munich Information center for Protein  
Sequences).  
[http://www.genoscope.cns.fr/externe/sequences/Banque\\_Projet\\_EF/EST](http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/EST)  
<http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis>.

## FEATURES

Location/Qualifiers  
source  
1..1059  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/ecotype="Col-0"  
/db\_xref="taxon:3702"  
/clone="GSLTFB65ZD02"  
/tissue type="Flowers and buds"  
/clone\_lib="Arabidopsis thaliana Flowers and buds Col-0"

## ORIGIN

Query Match 39.5%; Score 414.4; DB 5; Length 1059;  
Best Local Similarity 65.7%; Pred. No. 1.3e-112;  
Matches 661; Conservative 0; Mismatches 321; Indels 24; Gaps 3;  
Qy 1 AAGATGAAGATGAAGACAGTCTTGTATCATAACTGTATATCTTGGCCATTGGAAC 60  
Db |||||  
27 AAGAAGAAGATGAAGAATGGTTTGTATATCATAACTGTATATCTTACATATAGGAACA 86  
Qy 61 TGTGAGGCCCTCTAATGATCGTCTCTACTTCCAAATGTTGGCGAAAGGATCTGGTTT 120  
Db |||||  
87 TGTGAGGTCCTTGTGTTAACTCGTCTACTTCCAAATGCGGAAACGAATCTGGTT 146  
Qy 121 CCAAGCTCTCTCAAAACGGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 180  
Db |||||  
147 ATGAGCTTCCTATCAACCGCTGGTTTCCAAATCATCTCATCTCTCTCTCTCTCTCT 206  
Qy 181 CTCGCCGCTCGTGGTTG-----CCTTGAAGACAAGAACGACTC 220  
Db |||||  
207 CTCAGCGGTACGCGGACGCGAACCCCTAACACGCGGAAACAAAGCGGAAACAAA 266  
Qy 221 CATTTTTCCTCATGAACCTCTCTCTTTATCGCGCTATCGTTGTTGTTGTTGTTGTTG 280  
Db |||||  
267 GCTACTTCTCATGGAACCTCTCTGTTTATCGCTCCATGTCATAGGGTTGCTCACAG 326  
Qy 281 GATTGACAATTAACCTCTACTCTTACGCGGTAGTTATATCCCTGTTTCTACTGCTCTT 340  
Db |||||  
327 GACTTGACAACCTACTTATATTTCTTACGGAATAGCATATCTGCCAGTTTCAACTTCAT 386  
Qy 341 TGATCATCTCGCGCAATTAGGCTTCACTGCTCTCTTTCATTTTATGTTGTTGTTGTTG 400  
Db |||||  
387 TCATAATCGGAACCTCAACTAGCTTTCAACGCTCTCTTCGCTTCTTGAATATCAAGCAA 446  
Qy 401 AGTTTCAACCTTTTCACTATATAAAGCTAT-CGTTTGTCTCACTGTTGTTGTTGTTGTT 459  
Db |||||  
447 AGTTCACTCGCTTCTCCATAAACCGCGTACGTTTGTGAGGGTTGTTATCGGATCCTT 506  
Qy 460 GCCCTTAACCTCTGATAGTGAACAGCTTGTGAACAGACACAGGAATATGTTGTTGGG 519  
Db |||||  
507 TCTTTACACAGTAATGAAGAAAAACCGCTCTAAGGAAAAAATAAATAATTAATGGTTT 566  
Qy 520 TTTCATCATGACTCTTGGTCAGCTCTTCTCTATGGGTTTATATGTTGTTGTTGTTGTT 579  
Db |||||  
567 TTCTCGATGACTTTTTTTTTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 626  
Qy 580 TCTTACAAGAAATCTGTTGACGGAATACGCTATACGCTCGCTCGAGTTCCAGATGGTC 639  
Db |||||

Db 627 AATTAAAAAAGATCGTCAAGAAATCACTTTCCCACTGTGCTTGAGATTCAGATGGTC 686  
QY 640 TTATGCTTTGCTGCCACCTTGTGCTGCTCGTGGGATGCTAGCCGCTGCGCATTTCAAG 699  
Db 687 ATGTCCTTTCTTAATTCTTTCTTGTCATTTTCATTTTCATCTTTGGAGATTTAAG 746  
QY 700 GTGATAGCAGGAGCAAGCAAGATTTTAAAGCTTGGAGAGTCTTTG---TACTATGTTGGTG 756  
Db 747 GTGAAAGCAAGCAAGCAAGATAGTTCAAGATTTGGTGGATCATTTGTTTACTATTTCATTT 806  
QY 757 ATTGCTGTTACGGCCATAATCTGGCAAGCATTTTGTGGAGCTATTGCGTTGATCTTC 816  
Db 807 TTGTGATCACAGATATATAGCAAGTTCTTTCTTAGTGTCCAAAGGAATGTGTTT 866  
QY 817 TGTGATCGTCTCTGCTCTCTGGAATATATGCTCAGTCTGCTCTCCGGTGAAGGATGATC 876  
Db 867 TGTGATCATCACTAGCTCTCTGCTGTTCTGATAAGTGTCTGCTTCCGGTGACTGAATTT 926  
QY 877 TTGCGCGTCATTTGCTTTCAGGAGAAATTTTACGGCGGGAAGGTGTCTGTTGGCTCTC 936  
Db 927 TTGCGCGTGGTTTGTGTTTCCGGTTGTAGTTTTCAGTTAGTGTAAAGGTGTCTCTACTTCTT 986  
QY 937 TCCTCTCGGGATCACTCTCTTATTTCTATGGACAGTTTAAATCCG 982  
Db 987 TCATTTGGGTTTGTCTCATACATCACTAAGGTGTGTTTAAATCGG 1032

## RESULT 6

BH589732/c

LOCUS BH589732 688 bp DNA linear GSS 15-DEC-2001  
DEFINITION BOGES04TR BOGE Brassica oleracea genomic clone BOGES04, genomic survey sequence.

ACCESSION BH589732

VERSION BH589732.1 GI:17842184

KEYWORDS GSS

SOURCE Brassica oleracea

ORGANISM Brassica oleracea

REFERENCE 1 (bases 1 to 688)  
AUTHORS Town,C.D., Van Aken,S., Utterback,T., Koo,H. and Fraser,C.M.  
TITLE Whole genome shotgun sequencing of Brassica oleracea

JOURNAL Unpublished (2001)

COMMENT Other GSSs: BOGES04TF

Contact: Chris Town

TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA.

Tel: 301-838-3523

Fax: 301-838-0208

Email: cdtown@tigr.org

DNA is from a doubled haploid provided by Tom Osborn.

Seq primer: TR

Class: sheared ends.

Location/Qualifiers

1. .688

/organism="Brassica oleracea"

/mol\_type="genomic DNA"

/strain="TO100DH3"

/db\_xref="taxon:3712"

/clone\_lib="BOGES04"

/note="Vector: pHOS1; Site 1: BstXI; 2-3 kb sheared"

genomic DNA inserted into pHOS1 using BstXI linkers"

## FEATURES

source

## ORIGIN

Query Match 39.2%; Score 410.8; DB 8; Length 688;  
Best Local Similarity 77.3%; Pred. No. 1.4e-111;  
Matches 512; Conservative 0; Mismatches 147; Indels 3; Gaps 1;

QY 1 AAGATGAAGATGAAGACAGTCTTGTATCATATAAAGTGTATATCTTGGCCATTTGGAAC 60

Db 664 ATGATGAAGATGAAGCGGCTTCTTGTATCATATAAAGTGTATATATTTGGCGATAGGAAC 605

QY 61 TGTGGAGGCCCTCTAATGATGGTCTCTACTTTCCAAAATGTTGGCGAAAAGATCTGTTTT 120  
Db 604 TGTGGAGGCCCTCTAATGATGGTCTCTACTTTCCAAAATGTTGGCGAAAAGATCTGTTTT 545  
QY 121 CCAAGTCTCTTCAAAACGGTGTGTTGTCACATTTTCTTCCCTCTCTCTTATCTTTTC 180  
Db 544 TCAAGCTTCTCTCAAAACCTTCAGGTTGCCCATCATCTCTCTCTCTCTCTCTCTCTTC 485  
QY 181 CTCGCCGCTGCTGTTGCTTGAAGCAAGAAAC---GACTCCATTTTTCCTCATGAA 237  
Db 484 CTCACCGGTACCGTAAACAAGAACAGAAATATCAAGACTACACTCTTCTCTCATCAA 425  
QY 238 CTTCTCTCTTTATCGCGCTATCTGTTGTTGTTGCTCGTGGGATTTGACAAATACCTC 297  
Db 424 CTTCTCTGTTCTCTCGCTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 365  
QY 298 TACTCTTACGGGTTAGCTTATATCTGTTTCTACTGCGTCTTTGATCATCTCCGGCAA 357  
Db 364 TACGCTTACGGGTTAGCTTATCTCTCCCGTCTCCAGCTCTCTCTCTCATCTCTCTCAG 305  
QY 358 TTAGGCTTCACTGCTCTCTTTCATTTTATGTTGAAGCAAAAGTTTCACTTTTCACT 417  
Db 304 CTAGCTTACCGCTCTCTTTCGCTTCTTCATGGTGAAGCAAAAGTTTCACTTTTCACT 245  
QY 418 ATAAACGCTATCGTTTGTCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 477  
Db 244 ATAAACGCTGCTGTTTGTCTCACTCTCGGCGCGGAGCTTGTCTTTCACGGGACGT 185  
QY 478 GACAAGCTTGAACAGACACACAAAGAAATATGTTGTTGGGTTTCACTGATCTTTGTT 537  
Db 184 GACAAGCTTCCCAAGGAGACGCGAAGAGTATATAGTTGGGTTTGTATGACGGTCTGCT 125  
QY 538 CGAGCTCTCTATGGGTTTATATGTCACCTTGTGAGCTTCTTCAAGAAATCTGTT 597  
Db 124 CGAGCTTCTCTACGCTTGTGTTGCTGCTTGTGAGATGCAAGTTGTTGTTGTTGTTGTTG 65  
QY 598 CAGCGAAATCACTATACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 657  
Db 64 CAGAGATCAGCTACACCGCTTGTGTTGAGATGCAAGTTGTTGTTGTTGTTGTTGTTGTTG 5  
QY 658 TG 659  
Db 4 TG 3

## RESULT 7

BX836413/c

LOCUS

DEFINITION

BX836413 Arabidopsis thaliana Silique Col-0 Arabidopsis thaliana

cDNA clone GSLTSL532D06 3PRIM, mRNA sequence.

ACCESSION BX836413

VERSION BX836413.1

KEYWORDS

SOURCE

ORGANISM

Arabidopsis thaliana (thale cress)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

REFERENCE 1 (bases 1 to 463)

AUTHORS Castelli,V., Aury,J.M., Jallou,O., Wincker,P., Clepet,C.,

Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V.,

Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.

Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences:

A Combined Approach to Evaluate and Improve Arabidopsis Genome

Annotation

Unpublished (2004)

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

The sequences are based on single pass reads.

Life Technologies (a division of Invitrogen) members carried out

full-length libraries construction : Temple G.  
Genoscope members carried out sequencing and annotation : Castellani V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Sallanoubat M.  
URGV INRA : Clepet C., Caboche M.  
Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich Information center for Protein Sequences).  
[http://www.genoscope.cns.fr/externe/sequences/Banque\\_Projet\\_EP/EST](http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EP/EST)  
<http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis>.

## FEATURES

source  
Location/Qualifiers  
1. 463  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/ecotype="Col-0"  
/db\_xref="taxon:3702"  
/clone="GSLTSL53ZD06"  
/tissue\_type="Silique"  
/clone\_lib="Arabidopsis thaliana Silique Col-0"

## ORIGIN

Query Match 31.9%; Score 334.4; DB 5; Length 463;  
Best Local Similarity 97.4%; Pred. NO. 1e-88;  
Matches 372; Conservative 0; Mismatches 6; Indels 4; Gaps 3;  
  
Qy 672 GGGGATGCTAGCCGCTGGCGATTCAAGGTG-ATAGCAGGAGACGACGAGATTTTAAGC 730  
Db 463 GGGGATGCTAGCCGCTGGCGATTCAAGGTGATAGCAGGAGACGACGAGATTTTAAGC 404  
  
Qy 731 TTGAGAGCTCTTTGTACTATGTGGTGATGTGTTTACCGCCATATCTGGCAAGCATTTT 790  
Db 403 TTGAGAGCTCTTTGTACTATGTGGTGATGTGTTTACCGCCATATCTGGCAAGCATTTT 344  
  
Qy 791 TTGCGGAGCTATTGGGTGTGATCTTCTGTGCATCGTCTCTGGTCTCTGGAAATATGGTCA 850  
Db 343 TTGCGGAGCTATTGGGTGTGATCTTCTGTGCATCGTCTCTGGTCTCTGGAAATATGGTCA 284  
  
Qy 851 GTGCTCTGCTTCCGCTGACGCTGATCTT- -GCGCTCATTTTGTCTCCAGGAGAGTTTCA 908  
Db 283 GTGCTCTGCTTCCGCTGACGCTGATCTTCCGCCGTGCAATTTGCTTCCAGGAGAGTTTCA 224  
  
Qy 909 GCGGGGAAAGGTGTGCTTTGGCTCTCTCCCTCTGGGGATCAGTCTCTTATTTCTATGG 968  
Db 223 GCGGGGAAAGGTGTGCTTTGGCTCTCTCCATCTGGGATCAGTCTCTTATTTCTATGG 164  
  
Qy 969 ACAGGTTAAATCCGAGGAGAGACTAAGGCTCAGATACACAACTGTCTCAGCTTCAGT 1028  
Db 163 ACAGGTTAAATCCGAGGAGAGACTAAGGCTCAGATACACAACTGTCTCAGCTTCAGT 104  
  
Qy 1029 T-ACTGATTATGTAGCTTAAAA 1049  
Db 103 TCACGTGAATGTAGCTTAAAA 82

## RESULT 8

CK257466  
LOCUS CK257466 985 bp mRNA linear EST 30-JUL-2004  
DEFINITION ES741103 potato callus cDNA library, normalized and full-length  
Solanum tuberosum cDNA clone POC557 5' end, mRNA sequence.

## ACCESSION

CK257466

## VERSION

CK257466.1 GI:39814446

## KEYWORDS

## SOURCE

Solanum tuberosum (potato)  
Solanum tuberosum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; lamids; Solanales; Solanaceae; Solanum.

## REFERENCE

1 (bases 1 to 985)  
Buell.C.R., Hart.A., Zismann.V., Karamycheva,S.A. and Baker,B.  
Generation of ESTs from potato callus tissue  
Unpublished (2003)  
Contact: Robin Buell  
The Institute for Genomic Research

9712 Medical Center Dr, Rockville, MD 20850, USA  
Email: potato-array@tigr.org  
Clones can be requested from the University of Arizona Genomics  
Institute via <http://genome.arizona.edu/orders/>.  
Seq primer: ATT TAG GTG ACA CTA TAG.

## FEATURES

source  
Location/Qualifiers  
1. 985  
/organism="Solanum tuberosum"  
/mol\_type="mRNA"  
/cultivar="Kennebec"  
/db\_xref="taxon:4113"  
/clone="POCD557"  
/tissue\_type="callus"  
/lab\_host="DH10B-Tona"  
/clone\_lib="potato callus cDNA library, normalized and  
full-length"  
/notes="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;  
supplier: RNA was isolated from Solanum tuberosum var.  
Kennebec callus tissue grown on solid media."

## ORIGIN

Query Match 29.4%; Score 308.2; DB 7; Length 985;  
Best Local Similarity 58.4%; Pred. NO. 9.9e-81;  
Matches 560; Conservative 0; Mismatches 393; Indels 6; Gaps 1;  
  
Qy 3 GATGAAGATGAAGACAGTTCCTTGTATCATAACTGTATATCTTGGCCATTGGAAACTG 62  
Db 33 GAAATGAAGAAGCTTATCTTCTTAATTTTCAACATAAATATGCTTTCAATCGGCAATTG 92  
  
Qy 63 TGGAGGCCCTCTAATGATGCTCTCTACTTTCCAAAATGGTGGGAAAGAGTCTGGTTTCC 122  
Db 93 CGGTGGTCCCTTAATCTCTCGACTCTATTTCAITGCGGGCCACGAATTTGGATACC 152  
  
Qy 123 AAGTTCCTTCAACCCGTTGGTGTCCACTCATTTTCTTCCCTCTTCTCTTAICTTTCTCT 182  
Db 153 CAGCGTATTACAAACTGTTGGTTGTCCAATTAATTTCTCATCCCTCTAGCCATAGCTATTT 212  
  
Qy 183 CCGCGCTGCTGCTTGGCTTGAAGAACAGAAAGACCTCCATTTTCTCATGAACCTCC 242  
Db 213 CCAACGCCGGAATTC- - - - -AAGACCCGAGGCTAAATATGTCCTTTATCACACGGCA 266  
  
Qy 243 TCTCTTTATCCCGCTATCGTTGTGTTGTTGCTCGTGGGATTTGACAAATACCTCTACTC 302  
Db 267 AGAGTTTCATCGCATCTGCGGCGTGGGAATCATCTCGGTCTCGATGTTTACTTAAACTC 326  
  
Qy 303 TTACGGGTTAGCTTATATCCCTGTTTCTACTGGGTCTTTGATCATCTCCGGCAATTAGG 362  
Db 327 ATGGGGACCGCGAAATTTACCCGTTTCAACTTCGACATAATCAACCGCACTCAACTTGC 386  
  
Qy 363 CTTTCACTGCTCTCTTTTGCATTTTATGTTGAAGCAAAAGTTTCAACACCTTTCACTATAA 422  
Db 387 GTTCACTGCGCTTTTGTCTGTCTTATAGTTTAAACAGAAATTTGACAGCTATTTCGACGA 446  
  
Qy 423 CGCTATCGTTTGTCTCACTGCTGCTGCGGTAGTCTTTCCTTTCCTTAACTCTGATAGTACA 482  
Db 447 TTCGTGCTTTTGTCTTATCCGCGAGCTGCGACTTTTAGCTCTCCGGGGAACGGTGACCG 506  
  
Qy 483 GCTTGCAAAACGACACACAAAGAAATATGTTTGGGTTTCATCATGACTCTTGGTGAGC 542  
Db 507 GCGGCGCGCGAGTCCACGAAGGATTATATGTTGGGGTTTGTGATGACGTTTATTCGGTGC 566  
  
Qy 543 TCTTCTCTATGGGTTTATATTGCCACTTGTTCGAGCTTTTCTTACAAGAAATCTGCTCAGCG 602  
Db 567 GGTGTTATATGGATTTGATGTTGCGGTTAATGTAGTTGATTTATGAAGGGAAGCAAGC 626  
  
Qy 603 AATCACGTATACGCTCCGCTCGAGTTTCCAGATGGTCTTATGCTTTGCTGCCACTTGTGT 662  
Db 627 TGTACTTACACTACAGTATTGGAGATTTCAGATGGTTTGGGCAATTTCTCTACTCTGTTT 686  
  
Qy 663 CTGCTCGTGGGATGCTAGCCGTGGCGATTTTCAAGGTGATAGCAGGAGAACGAAGA 722  
Db 687 TTGCACCTATTGGAATGGTTTATCAACAGAGGATTTTCCAGGCGATTTCCAGGGAAGCAAGTCA 746

```
QY 723 TTTTAAAGCTGGAGAGTCTTTGTACTATGTGGTGAATGTGGTTCACGGCCATAATCTGGCA 782
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
747 ATATGCAATTGGAGAAATCTAAGTATTATATATGGTACTAGTATGGTGTGCCATTATTTGGCA 806
QY 783 AGCAATTTTGTGGGAGCTATTGGGTGTATCTTCTGTGATCGTCTCTGGTCTCTGGAAT 842
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
807 ATTCCGGCTCTCTGGGCTAGTGGAGTTATTTTTTATTTCTTCTTACTCTCTGGAAT 866
QY 843 TATGGTCAGTCTCTCTCTCCGGTGACGGTGAATCTTGGGCGTCAATTTGCTTCCAGGAGAA 902
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
867 TATAGGGCTTTTACTTCTCTGTACCCGAGTTTGGCTGTAAATTTTGTTCACGAGAAA 926
QY 903 GTTTCAGGGGGGAAAGGTGTGCTCTTGGCTCTCTCCCTCTCGGGATCAGTCTCTTATT 961
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
927 ATTTCAAGCCGAAAAGGAGTAGTATTTTCTAGCTCTATGGGATTTGTTTCATATT 985

RESULT 9
CK243751 986 bp mRNA linear EST 30-JUL-2004
LOCUS EST727388 potato callus cDNA library, normalized and full-length
DEFINITION Solanum tuberosum cDNA clone POCAS40 5' end, mRNA sequence.
ACCESSION CK243751
VERSION 1
KEYWORDS 1 (bases 1 to 986)
SOURCE Generation of ESTs from potato callus tissue
ORGANISM Unpublished (2003)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum.
REFERENCE Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A. and Baker, B.
AUTHORS Generation of ESTs from potato callus tissue
TITLE Unpublished (2003)
JOURNAL
COMMENT Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@tigr.org
Clones can be requested from the University of Arizona Genomics
Institute via http://genome.arizona.edu/orders/.
Seq primer: ATT TAG GTG ACA CTA TAG.
FEATURES
source
Location/Qualifiers
1..986
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="POCAS40"
/tissue_type="callus"
/lab_host="DH10B-TonA"
/clone_lib="potato callus cDNA library, normalized and
full-length"
/notes="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;
supplier: RNA was isolated from Solanum tuberosum var.
Kennebec callus tissue grown on solid media."
```

## ORIGIN

```
Query Match 29.2%; Score 306.2; DB 7; Length 986;
Best Local Similarity 58.4%; Pred. No. 3.9e-80;
Matches 575; Conservative 0; Mismatches 403; Indels 7; Gaps 2;

QY 3 GATGAAGATGAAGACAGTCTTGTATATCAATACTGTATATTCTTGGCCATTGGAACTG 62
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
4 GAAATGAAGAGCTTATCTTCTTAATTTCAACATAATTATGCTTACATCGGCAATTG 63
QY 63 TGGAGGCCCTCTAATGATGGCTCTTACTTCCAAATGTGGCGAAGATCTGGTTCC 122
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
64 CGGTGGTCTTAAATCTCTCGACTTATTTCAATTCATTCATGGCGCCACGAATTTGGATACC 123
QY 123 AAGCTTCTTCAACCGTTGGTGTCCACTATTTTCTCCCTCTCTCTTATCTTCT 182
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
124 CAGCGTATTACAAACGTTGGTGTGCCAATTTCTATCCCTCTAGCCATAGCCTATTT 183
```

```
QY 183 CCGCGCTGCTGCTTGGCTTTGAAGAAACAGAACTCCATCTTTTCTCATGAACCTCC 242
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
184 CCAACGCCGGAATAATTC-----AAGGACCCGAGGCTAAATCGTCTTTATCACGGCA 237
QY 243 TCTCTTTATCGCGCTATCGTTGTTGTTGCTGCTGGGATTTGACAAATTAACCTCTACTC 302
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
238 AGAGTTTCATCGCATCTGCGCGCTCGGAATCATCTCGGTCTCGATGGTTACTTAACTC 297
QY 303 TTACGGGTTAGCTTATATCCCTGTTTCTACTCGCTCTTTTGATCATCTCCGCGCAATTAG 362
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
298 ATGGGAGCCCGGAAATTAACCGTTTCACTTCGACATAATCAAGCGACTCAACTTGC 357
QY 363 CTTCATCTCTCTTTTGCATTTTATGGTGAAGCAAAAGTTTCAACCTTTCACTATAAA 422
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
358 GTTCACTGCGCTTTTGTCTTATAGTTAAACAGAAATTCACAGCGTATTCGACGAA 417
QY 423 CGCTATCGTTTGTCTCACTGGTGGTCCGTAGTCTTGGCCCTTAACCTGATAGTGACAA 482
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
418 TTCCGCTGCTTTTGTCTTATCGCGGAGCTGCGACTTTAGCTCTCCGGCGAACGGTACCG 477
QY 483 GCTTGCAAAACGAGACACACAGGAATATGTTCTGGGTTTCATCATGACTCTTTGGTGCAGC 542
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
478 GCGGCGCGGAGTCCACGAGGATATATGTTGGGTTTGTGATGACGTTTTATCGGTGC 537
QY 543 TCTTCTCTATGGGTTTATATGCCACTTGTCCAGCTTTCTTACAAGAAATCTGGTCAGCG 602
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
538 GGTGTTATATGATGATGTTGCTTAAATGAGTTGATTTATATGAAGCGAAGCAAGC 597
QY 603 ATACAGTATAGCTCGGCTCGATTCAGATGCTTATGCTTTGCTGGCCACTTGTGT 662
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
598 TGTTACTTTACACTACAGTATTTGGAGATTCAGATGTTTTTGGGCATTTCTGCTACTGTTT 657
QY 663 CTGCTCTGTTGGGATGCTAGCGCTGAGCTTTTCAAGTGTAGCAGAGAGCAAGAGCA 722
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
658 TTGCACTATTTGAATGATATCAACAGGATTTCCAGGCGATTCAGGGAAGCAAGTCA 717
QY 723 TTTTAAAGCTTGGAGAGTCTTTGTACTATGTGGTGAATGTTTTCAGCGCCATAATCTGGCA 782
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
718 ATATGCAATTTGAGAATCTAAGTATTATATGTTAGTATGTTGCTGTCATTTTGGCA 777
QY 793 AGCATTTTGTGGAGCTATTTGGTGTGATCTTCTGTGATCGTCTCTGCTCTGGAAT 842
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
778 ATTCGCGCTCTCTGGGTTAGTTGGAGTTATTTTATTTCTTCTCTTACTCTCTGGAAT 837
QY 843 TATGTCAGTCTCTGCTTCCGCTGACGCTGATCTTGGCCGTCATTGCTTCCAGGAGAA 902
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
838 TATAGCGCTTTTCTTCTCTGTTACCAAGTTTGGCTGTAAATTTGTTCCAGAGAAA 897
QY 903 GTTTCAGGCGGGAAGGTGTCGCTTTGGCTCTCTCCCTCT--GGGATCAGTCTCTTATT 961
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
898 ATTTCAAGCCGAAAAGGAGTAGCTATTTTCTAGCTCTATGGGGAATTTGTTTCATATT 957
QY 962 TCTATGACAGGTTAAATCCGAGGA 986
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
958 TTTATGTTGAATATGAACAAGGGGA 982

RESULT 10
BG439880 680 bp mRNA linear EST 15-MAR-2001
LOCUS GA_Ea005C03f Gossypium arboreum 7-10 dpa fiber library Gossypium
DEFINITION arboreum cDNA clone GA_Ea005C03f, mRNA sequence.
ACCESSION BG439880
VERSION 1
KEYWORDS 1 (bases 1 to 680)
SOURCE EST.
ORGANISM Gossypium arboreum
Gossypium arboreum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
REFERENCE Wing, R.A., Frisch, D., Yu, Y., Main, D., Rambo, T., Simmons, J.,
AUTHORS Henry, D., Wood, T.C., Leslie, A. and Wilkins, T.A.
```

TITLE An integrated analysis of the genetics, development, and evolution of the cotton fiber

JOURNAL Unpublished (2000)

COMMENT Contact: Wing RA  
Clemson University Genomics Institute  
100 Jordan Hall, Clemson, SC 29634, USA  
Tel: 864 656 7288  
Fax: 864 656 4293  
Email: rwing@clemson.edu  
Seq primer: TAAACGACTCACTATAGGG  
High quality sequence stop: 621.  
Location/Qualifiers  
1. .680  
/organism="Gossypium arboreum"  
/mol\_type="mRNA"  
/strain="AKA"  
/cultivar="8400"  
/db\_xref="taxon:29729"  
/clone="GA\_Ea005C03f"  
/tissue\_type="Fibers isolated from bolls harvested 7-10 dpa"  
/lab\_host="E. coli"  
/clone\_lib="Gossypium arboreum 7-10 dpa fiber library"  
/note="Vector: pBK-CMV; Site\_1: EcoRI; Site\_2: XhoI"

FEATURES  
source

ORIGIN

Query Match 29.0%; Score 304.6; DB 4; Length 680;  
Best Local Similarity 65.5%; Pred. No. 1e-79;  
Matches 445; Conservative 0; Mismatches 234; Indels 0; Gaps 0;

QY 290 ATTACTCTACTCTTACGGGTAGCTTATATCCCTGTTTCTACTCGCTTTGATCATCT 349  
DB 1 ACTACTTCTACTCTTATCTCGTCTCCCGTTTCGACTTCTCTTTGATCATCG 60

QY 350 CGCGCAATTAGGCTTCACTGCTCTCTTTGCAATTTTATGGTGAAGCAAAAGTTCAAC 409  
DB 61 GGTGCAAGTGGCTTTTCAAGCGGGGTTTGTCTTCTGTTGGTGAACAAAGTTCACTG 120

QY 410 CTTTCACTATAAACCTATCGTTTGTCTCACTGCTGGTGGCGGTAGTCTTCCCTTACT 469  
DB 121 CCTACTCATAAACCGCGTGTGTGTGACATAGGGGCTGGCGTTCTGGCTTTGCATT 180

QY 470 CTGATAGTGACAAGCTTGCAACACGACACACAAAGAAATATGTTGGGTTCATCATGA 529  
DB 181 CGAGCAGCGACGGGCTGAGATGAATCCACAAAGAAATATTTTGGGTTCTGTAATGA 240

QY 530 CTCCTGGTGCAGCTCTTCTATAGGGTTTATTTGCCACTTGTGAGGCTTCTTACAAGA 589  
DB 241 CTTTACGCTGCAGCGCTTTGTACGGATTTGTATTCCTTTGGTGAACCTAACGTAAGA 300

QY 590 AATCTGGTCAGGAATCAGTATACGCTCGCGCTGAGTTCAGATGCTTATGCTTTG 649  
DB 301 AGGCAAGCAAGAGATCAGTACGCGCTTGTGTGAGGATTCAGCTGGTGTATGTTGG 360

QY 650 CTGCCACTTGTCTGCTCGCTGGGGATGCTAGCGGCTGGCGATTCAAGGTGATAGCAG 709  
DB 361 TTGCTACTGGCTTTTGACCGGTGGATGCTGGTCAACAATGATTTCAAGGTGATCGAA 420

QY 710 GAGAAGCAAGAGATTTTAAGCTTGGAGAGTCTTTGTACTATGTGGTGAATGTGTTACGG 769  
DB 421 AGGAAGCAAGGAATTCGAGCTAGGAGAAACAAATACTATGTGGTGTGATTTTACTG 480

QY 770 CCATAATCTGCAAGCAATTTTGTGGGAGCTATTGGGTGATCTTCTGTGCATCGTCTC 829  
DB 481 CAATTAATAATCAATGTTTCTTCTGGGAGCTATTGGAGTAGTATTTGTGGCTCATCA 540

QY 830 TGGTCTCTGGAAATATGCTCAGTGTCTGCTTCCGGTGACGGTGAATCTTGGCCGCTATT 889  
DB 541 TGCTATCGGGTGTAGTAGCGGTTCTATTACAGTTTACAGATTTTGGCAATATTTT 600

QY 890 GCTTCCAGAGAAAGTTTCAAGCGGGGAAAGGTGCTGCTTCTCCCTCTGCGGAT 949  
|||||

DB 601 CTTTCAATGAGAGCTTTTCAAGCCGAGAAAGGTGTGTCTCTTGTCTCTCTCGCTGTGAGCT 660

QY 950 CAGTCTCTTATTCTATG 968  
DB 661 TCGATAACTACTCTTTATGG 679

RESULT 11  
CK247918  
LOCUS CK247918 998 bp mRNA linear EST 30-JUL-2004  
DEFINITION EST731555 potato callus cDNA library, normalized and full-length  
Solanum tuberosum cDNA clone POXAX15 5' end, mRNA sequence.  
ACCESSION CK247918  
VERSION CK247918.1 GI:39797543  
KEYWORDS EST  
SOURCE Solanum tuberosum (potato)  
ORGANISM Solanum tuberosum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; lamids; Solanales; Solanaceae; Solanum.  
REFERENCE 1 (bases 1 to 998)  
AUTHORS Buell,C.R., Hart,A., Zismann,V., Karamycheva,S.A. and Baker,B.  
TITLE Generation of ESTs from potato callus tissue  
JOURNAL Unpublished (2003)  
COMMENT Other ESTs: EST731556  
Contact: Robin Buell  
The Institute for Genomic Research  
9712 Medical Center Dr, Rockville, MD 20850, USA  
Email: potato-array@tigr.org  
Clones can be requested from the University of Arizona Genomics  
Institute via http://genome.arizona.edu/orders/  
Seq primer: ATT TAG GTG ACA CTA TAG.  
Location/Qualifiers  
1. .998  
/organism="Solanum tuberosum"  
/mol\_type="mRNA"  
/cultivar="Kennebec"  
/db\_xref="taxon:4113"  
/clone="POXAX15"  
/tissue\_type="callus"  
/lab\_host="DH10B-Tona"  
/clone\_lib="potato callus cDNA library, normalized and full-length"  
/note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;  
supplier: RNA was isolated from Solanum tuberosum var.  
Kennebec callus tissue grown on solid media."

ORIGIN

Query Match 28.8%; Score 302.2; DB 7; Length 998;  
Best Local Similarity 58.3%; Pred. No. 6.3e-79;  
Matches 551; Conservative 0; Mismatches 388; Indels 6; Gaps 1;

QY 3 GATGAAGATGAAGACAGTCTTGTATATCATAACTGTATATTTCTGGCCATTGGAAACTG 62  
DB 25 GAAATGAAGAAGCTTATCTTCTTAATTTTCAACATAAATATGCTTTCAATCGCAATTG 84

QY 63 TGGAGGCCCTCTAATGATGCTCTACTTCCAAAATGGTGGCAAGGATCTGTTTCC 122  
DB 85 CGGTGGTCTTAACTCTCTGACTCTATTTCAATCATGGCGCCACGAATTTGATACC 144

QY 123 AAGCTTCCTTCAAAACCGTGTGTGTCACCTCATTTTCTCCCTCTCTCTTATCTTTCT 182  
DB 145 CAGCGTATTACAACTGTGTGTGTCCAATTTATCTCATCCCTCTAGCCATAGCTATTT 204

QY 183 CGCGCGTCTGCTTGGCTTGAAGACAAAGAACGCTCCATTTTCTCTCATGAACCTCC 242  
DB 205 CCAACGCCGGAATAATTC-----AAGGACCCGAGGCTAAATACTCTTTATCACACGCA 258

QY 243 TCTCTTATCGCGCTATCGTTGTGTTGTTGCTCGTGGGATTTGACAATTAACCTCTACT 302  
DB 259 AGAGTTCAATCGACTCTCCCGGCTCGGAATCATCTGTCGGTCTCGATGTTACTTAACTC 318

QY 303 TTACGGGTAGCTTATATCCCTGTTTCTACTGCGTCTTTTGATCATCTCCGCGCAATTAGG 362



```
Db 319 ATGGGACCGCGGAAATACCCGTTTCACTTCGACATAATCAACGCGACTCAACTGC 378
QY 363 CTTCACTGCTCTCTTGGCAATTTTATGGTGAAGCAAAAGTTTCACACCTTTTCACTATAA 422
Db 379 GTTCACTGCGCTTTTGTCTGCTTATAGTAAACAGAAATTCAGACGATTCGACGAA 438
QY 423 CGCTATCGTTTCTCACTGCTGCTGCGGAGTTCCTGGCCCTTAACTCTGATAGTACAA 482
Db 439 TCCGTCGTTTCTTATCGCGGAGCTGCGACTTTAGCTCTCCGGCGAAGCTGACCG 498
QY 483 GCTTGCACAGCAGACACACAGGAATATGTTGGGTTTCATCATCACTCTTGGTCACG 542
Db 499 GCCGGCGCGAGTCACAGGAATATATGTTGGGGTTTGTGATCAGCTTATCGGTCG 558
QY 543 TCTCTCTATGGGTTTATATGTCACCTGTCGAGCTTTCTTACAAGAAATCTGGTCAGCG 602
Db 559 GGTGTTATATGGAATGATGTTGCCGTTAATGAGTTGATTTATATGAAGCGCAAGC 618
QY 603 AATCAGCTATACGCTCGCGCTCGAGTTCAGATGGTCTTATGCTTTGCTGCCACTTGTG 662
Db 619 TGTACTTACACTACAGTATTCAGATTCAGATGGTTTGGGCAATTTCTGCTACTGTTT 678
QY 663 CTGCTCGTGGGATGCTAGCCGCTCGGATTTCAAGGTGATAGAGGAGCAAGAGA 722
Db 679 TTGCACATTTGGAATGGTTATCAACAGGAATTTCCAGGCGATTCGAAGGCAAGTCA 738
QY 723 TTTTAAAGCTTGGAGAGCTTTTGTACTATGTTGGTATGTTGTTCCAGCCATATCGGCA 782
Db 739 ATATGCAATGGAGAACTAAGTATTTATATGTTACTAGTATGTTGGCAATTTTGGCA 798
QY 783 AGCAATTTTGTGGGAGCTATTGGGTTGATCTCTGTCATCGTCTCTGCTCTGGAAT 842
Db 799 ATTCCGGCTCTCGGGTTAGTTGGAGTTATTTTATTTCTCTCTCTCTCTGGAAT 858
QY 843 TATGTCAGTCTCTGCTCCGCTGACGCTGATCTTTGGCGGCAATTTGCTTCAGAGAA 902
Db 859 TATAGCGCTCTTTTACTTCTCTGTTACCGAAGTTTGGGCTGTAATTTGTTCCAGAAA 918
QY 903 GTTTTACGGCGGGAAGGTGTCGCTTTGGCTCTCTCCCTCTGGG 947
Db 919 ATTTCAAGCGGAAAAGGAGTAGCTATTTTCTAGCTCTATGGG 963
```

## RESULT 12

CK251896 1013 bp mRNA linear EST 30-JUL-2004  
LOCUS EST35533 potato callus cDNA library, normalized and full-length  
DEFINITION Solanum tuberosum cDNA clone POCBR03 5' end, mRNA sequence.

CK251896

CK251896.1 GI:39805365

EST.

Solanum tuberosum (potato)

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; lamids; Solanales; Solanaceae; Solanum.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A. and Baker, B.  
Generation of ESTs from potato callus tissue  
Contact: Robin Buell  
The Institute for Genomic Research  
9712 Medical Center Dr, Rockville, MD 20850, USA  
Email: potato-array@tigr.org  
Clones can be requested from the University of Arizona Genomics  
Institute via <http://genome.arizona.edu/orders/>.  
Seq primer: ATT TAG GTG ACA CTA TAG.

FEATURES

source

1..1013

/organism="Solanum tuberosum"

/mol\_type="mRNA"

/cultivar="Kennebec"

/db\_xref="taxon:4113"  
/clone="POCBR03"  
/tissue\_type="callus"  
/lab\_host="DH10B-Tona"  
/clone\_lib="potato callus cDNA library, normalized and  
full-length"  
/note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;  
supplier: RNA was isolated from Solanum tuberosum var.  
Kennebec callus tissue grown on solid media."

## ORIGIN

Query Match 28.7%; Score 301.4; DB 7; Length 1013;  
Best Local Similarity 58.2%; Pred. No. 1.1e-78;  
Matches 569; Conservative 0; Mismatches 401; Indels 7; Gaps 2;  
QY 3 GATGAAGATGAAGACAGTTCCTTGTATCATTAACCTGTATATCTTGGCCATTTGGAACG 62  
Db 40 GAAATGAAGAAGCTTATCTTCTTAATTTTCAACATAATATATGCTTTCAATCGGCAATG 99  
QY 63 TGGAGGCCCTCTAATATGATCGCTCTCTACTTCCAAATGGTGGCGAAAGGATCTGGTTCC 122  
Db 100 CGGTGTCCTTTAATCTCTGACTCTATTTTCAATCATGGCGCCACGAATTTGGATACC 159  
QY 123 AAGCTTCTTCAAACCGTTGGTTGCCACTCAATTTTCTTCCCTCTTCTCTTATCTTCT 182  
Db 160 CAGCGTATTAACAACATGTTGGTTGTCCAATTTATCTCATCCCTCTAGCCATAGCCTAT 219  
QY 183 CGCGCGTCTGTCGCTTGAAGACAGAAACGACTCCATTTTCTCTCATGAAACCTCC 242  
Db 220 CCAACGCCGGAATAATTC-----AAGGACCCGAGGCTAAATCGTCTTTATCACAGGCA 273  
QY 243 TCTCTTTATCGCGCTATCTGTTTGTGTTTGTCTGGGATTTGACAAATTAACCTCTACTC 302  
Db 274 AGAGTTTCATCGCATCTGCGCGGCTCGGAATCATCTGCTCTCGATGTTTACTTAACTC 333  
QY 303 TTACGGGTTAGCTTATATCCCTGTTTCTACTCGGCTTTTGTGATCATCTCCGCGCAATTAG 362  
Db 334 ATGGGGACCCCGGAAATTAACCGTTTCAACTTCGACACTTAATCAACGCGACTCAACTTCG 393  
QY 363 CTTCACTGCTCTCTTTTGGCAATTTTATGTTGAAGCAAAAGTTTCAACCTTTCACTATAA 422  
Db 394 GTTCACTGCGCTTTTGTGCTTATAGTTAAACAGAAATTCAGAGCTATTCGACGAA 453  
QY 423 CGCTATCGTTTGTCTCACTGCTGCGGCTAGTCTCTGCTTAACTCTGATAGTACAA 482  
Db 454 TTCCGCTGTTTGTCTTATCGCGGAGCTGCGACTTTAGCTCTCCGGCGAACGCTGACCG 513  
QY 483 GCTTCAAAACGAGACACACAGGAATATGTTGTTGGGTTTCACTGACTCTTGGTGACGC 542  
Db 514 GCCGCGCGCGAGTCCAGGAAGTATATGTTGGGTTTGTGATGACGCTTTATCGGTGC 573  
QY 543 TCTTCTCTATGGGTTTATATTTGCCACTTGTCCGAGCTTCTTACAGAAATCTGGTCAGCG 602  
Db 574 GGTGTTATATGGAATGATGTTGCCGTTAATTTGAGTTGATTTATATGAAGCGAAGCAAGC 633  
QY 603 AATCAGCTATACGCTCGCGCTCGAGTTCAGATGCTTATGCTTTGCTGCCACTTGTCT 662  
Db 634 TGTACTTACACTACAGTATTCAGATTCAGATGTTTGGCAATTTCTGCTACTGTTT 693  
QY 663 CTGCTCTGTTGGGATGCTAGCGCTGGCGATTTTCAAGGTGATAGCAGGAAGCAAGAGA 722  
Db 694 TTGCACATTTGGAATGGTTATCAACAGGAATTTCCAGGCGATTCGAAGGCAAGCAAGTCA 753  
QY 723 TTTTAAAGCTTCGAGAGTCTTTGTACTATGTTGTTGTTTTCAGCGCCATATCTGGCA 782  
Db 754 ATATGCAATTTGGAATCTAAGTATATATGTTACTAGTATGTTGTTGCTTATTTGGCA 813  
QY 783 AGCAATTTTGTGGGAGCTATTGGGTTGATCTTCTGTCATCGTCTCTGCTCTCTGGAAT 842  
Db 814 ATTCCGCTCTCGGGTTAGTTGGAGTTATTTTATTTCTTCTCTTACTCTCTGGAAT 873  
QY 843 TATGTCAGTGTCTGCTTCCGCTGACCGGTGATCTTGGCCGTCATTTGCTTCCA-GGAGA 901



```

source
1. 979
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="POC8R86"
./tissue_type="callus"
/lab_host="DH10B-TonA"
/clone_lib="potato callus cdna library, normalized and
full_length"
/note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;
supplier: RNA was isolated from Solanum tuberosum var.
Kennebec callus tissue grown on solid media."

```

Query Match	27.7%	Score 290.6	DB 7	Length 979
Best Local Similarity	58.2%	Pred. No. 1.9e-75		
Matches 551	Conservative 0	Mismatches 389	Indels 7	Gaps 2
Qy	3	GATCAAGATGAAGACAGTCTCTTGTAATCATATAAATCTGTATATCTTGGCCATTTGAAACTG	62	
Db	39	GAAATGAAGAAGCTTATCTCTCCATAATTTTCAACATAATATATGCTTTCAATCGGCAATTG	98	
Qy	63	TGAGGGCCCTTAATGATGCGTCTCTACTTCCAAAATGGTGGCGAAGAGATCTGTTTCC	122	
Db	99	CGTGGTCTTAAATCATCTCGACTCTATTTCATTATGCGCGCCCAACGAATTTGGATACC	158	
Qy	123	AAGCTTCTCAAAACGGTTGGTGTGCACCTCAATTTTTCCTCTTCTTTATCTTTTCC	182	
Db	159	CAGCGTATTACAAACTGTGGTGTGCCAATTAATTTCTCATCCTCTAGCCATAGCCTATTT	218	
Qy	183	CCGCGGTCGTGTGCTTGAAGAACAGAAACGACTCCATTTTCTCATGAACCTCC	242	
Db	219	CCAAACGGCGAAATTC-----AAGGACCGAGGCTAAATTCGCTTTATCACACGGCA	272	
Qy	243	TCTCTTTATCGCCGCTATCGTTGTGGTTTGTCTCGTGGGATTTGACAATTAACCTCTACTC	302	
Db	273	AGAGTTTCATCGCATCTCGCGGTCGGAATCATCGTGGTCTCGATGGTTACTTTAACTC	332	
Qy	303	TTACGGGTTAGCTTATATCCCTGTGTTCTTACGCGTCTTTGATCATCTCCGCGCAATTAGG	362	
Db	333	ATGGGACCGCGGAAATTTACCGGTTTCAACTTCGACTCTAATCAACGCGACTCAACTTGC	392	
Qy	363	CTTCACATGCTCTTTTGCATTTTATGCGTGAAGCAAAAGTTACACCTTTCACATATAA	422	
Db	393	GTTTCACTGCGCTTTTGTCTGTCTTAGTTTAAACAGAAATTGACAGCGTATTTCGACGA	452	
Qy	423	CGCTATCGTTTTGTCTCACTGCGTGGTCCGTAGTCCCTTGCCCTTAACCTCTGATAGTGACA	482	
Db	453	TTCCGTGTTTTGCTTATTCGCGGAGCTCGGACTTTAGCTCTCCGGGCGAACGCTGACCG	512	
Qy	483	GCTTGCAACGAGACACACAAAGGAATATGTTGTGGGTTTCAATGACTCTTTGGTGCAGC	542	
Db	513	GCGCGCGCGGAGTCCACGAAGGAATATATGTTGGGGTTTGTGATGACGTTTATCGGTGC	572	
Qy	543	TCTTCTCTATGGGTTTATATTTGCCACTTGTCCGACTTTCTTACAAAGAAATCTGGTTCAGCG	602	
Db	573	GGTGTTATATGGAATGATGTTGCCGTTAATTAAGTTGATTTATGATGAAGCGAAGCAAGC	632	
Qy	603	AATCACGATACCGTCGCGCTCGAGTTCAGATGGTCTTATGCTTTTCTGCTCACTTGTGT	662	
Db	633	TGTTACTTACACTACAGTATTTGGAGATTCAGATGGTTTGGGCAATTTCTGCTACTGTTTT	692	
Qy	663	CTGCTCGTGGGATGCTAGCCGCTGGGATTTCAAGTGTATAGCAGAGAGAACGAAGA	722	
Db	693	TTGCACCTATTGGAAATGATTAACAAGGAATTTCCAGGCGGATTCGAAGGGAAGCAAGTCA	752	
Qy	723	TTTTAAAGCTGGAGAGTCTTTGTACTATGTGGTGAATTTGTGTTACGGGCCATAATCTGC	782	
Db	753	ATATGCAATTTGGAGAACTAAGTATTTATATGGTACTAGTATGGTGTGCCATTATTGGCA	812	
Qy	783	AGCAATTTTTTGTGGAGCTATTTGGTTGAATCTTCTGTGCAATCGGCTCTGTGCTCTGGAAT	842	

Db	813	ATTCGCGCTCTCGGGGTAGTTGGAGATTATTTTTTATCTTCTCTTATCTCTCTGGAAT	8752
Qy	843	TATGGTCAGTAGCTCTGCTTCCGGGTGACGGGTGATCTTGGCCGTCAATTTGGCTTCACAGGAA	902
Db	873	TATAGGCGCTTTTTTACTTCTGTTTACGAAAGTTTGCTGTAATTT-TTTTCACGAAA	931
Qy	903	GTTTACGCGGGGAAAGGTCGCTTTGGCTCTCTCCCTCTGGGGAT	949
Db	932	ATTTCAAGCCGAAAAAGCAGTAGCTATTTTCTAGCTCTATGGGGAT	978

  

RESULT	15	
CK248243		
LOCUS	CK248243	988 bp mRNA linear EST 30-JUL-2004
DEFINITION	EST731880 potato callus cdna library, normalized and full-length	
	Solanum tuberosum cdna clone POC4244 5' end, mRNA sequence.	
ACCESSION	CK248243	
VERSION	CK248243.1	GI:39798187
KEYWORDS	EST.	
SOURCE	Solanum tuberosum (potato)	
ORGANISM	Solanum tuberosum	
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamids; Solanales; Solanaceae; Solanum.	
AUTHORS	1 (bases 1 to 988)	
TITLE	Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A. and Baker, B.	
JOURNAL	Generation of ESTs from potato callus tissue	
COMMENT	Unpublished (2003)	
	Contact: Robin Buell	
	The Institute for Genomic Research	
	9712 Medical Center Dr. Rockville, MD 20850, USA	
	Email: potato-array@tigr.org	
	Clones can be requested from the University of Arizona Genomics	
	Institute via http://genome.arizona.edu/orders/	
	Seq primer: ATT TAG GTG ACA CTA TAG.	

## FEATURES

```

1. 388
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="POCA24"
/tissue_type="callus"
/lab_host="DH103-Tona"
/clone_lib="potato callus cDNA library, normalized and
full-length"
/note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;
supplier: RNA was isolated from Solanum tuberosum var.
Kennebec callus tissue grown on solid media."

```

## ORIGIN

Query Match	27.5%	Score 288.4;	DB 7;	Length 988;
Best Local Similarity	57.8%;	Pred. No. 8.7e-75;		
Matches 554;	Conservative	0;	Mismatches 396;	Indels 8;
				Gaps 2;

Qy	3	GATGAGATGAAGACAGTCTCTTGTAATCATCAAACTGTATATTCTTGCCCAATTCGAAACTG	62
Db	39	GAAATGAAGAGCTATCTTCTCTAATTTTCAACATAAATTATGCTTTCAATCGCAATTG	98
Qy	63	TGGAGGCCCTCTAATGATGCGTCTCTACTTCCAAAATGTGGGAAAGGATCTGGTTTCC	122
Db	99	CGGTGGTCCTTTAATCTCTCGACTTATTTCAATTCATGCGGCCAACGAATTTGGATACC	158
Qy	123	AAGCTTCCTTCAAACCGTGTGGTTGTGCACATCTTTTCTCCCTCTCTCTCTATCTTTCCCT	182
Db	159	CAGGTAATTACAACATGTTGGTTGTCCAAATTATCTCATCCCTCTAGCCATAGCCTATTT	218
Qy	183	CCGCGTCGTCGTGGCTTGAAGAACACAGAACGACTCCATTTTTTCTCATGAACCTCC	242
Db	219	CCAACGCCGGAAATTC-----AAGAACCGAGGCTAAATCGTCTTTATCATCACGGCA	272
Qy	243	TCCTCTTATCGCGCTATCGTTCTTGTTTGGTCGTGGGATTTGACAAATTTACCTCTACTC	302

Db 273 AGAGTTTCATCGCATCTGCGGCGTCGGAAATCATCGTCGGTCTCGATGGTTACTTTAAACTC 332  
Qy 303 TTACGGGTTAGCTTATATCCCTGTTTCTACTGCGTCTTTGATCATCTCCGGGAAATTAGG 362  
Db 333 ATGGGACCCCGGAAATTTACCGTTTCAACTTCGACTCTAATCAACGCGACTCAACTGC 392  
Qy 363 CTTCACGTCTCTCTTTGCAATTTTATGGTGAAGCAAAAGTTACACCTTTTCACTATAAA 422  
Db 393 GTTCACTGCGCTTTTGTGTGCTTATAGTTTAAACAGAAATTTGACAGGTATTCGACGAA 452  
Qy 423 CGTATCGTTTGTCTCACTGGTGTGCGGTAGTCTTGGCCCTTAACTCTGATAGTGACAA 482  
Db 453 TTCGTCGTTTGTCTTATCGCGGAGCTGCGACTTTAGCTCTCCGGGGGAAAGGTGACCG 512  
Qy 483 GCTTGCAAAACAGACACACAAAGGAATATGTTGTTGGGTTCATCATGACTCTTTGGTGCAGC 542  
Db 513 GCCGCGCGCGAGTCCACGAAGGATATATGTTGGGGTTTGTGATGACGTTTATCGGTGC 572  
Qy 543 TCTTCTCTATGGGTTTATATGCCACTTGTGCGACTTCTTACAGAAATCTGGTCAGCG 602  
Db 573 GGTGTTATGGAATGATGTTGCCGTTAAATTGAGTTGATTTATATGAAGGCGAAGCAAGC 632  
Qy 603 AATCAGGTATACGCTCGCGCTCGAGTTCAGATGGTCTTATGCTTTGCTGCCACTTGTGT 662  
Db 633 TGTACTTACACTACAGTATTTGGAGATTCAGATGTTTGGGCAATTTCTGCTACTGTTT 692  
Qy 663 CTGCTCGTGGGATGCTAGCCGCTGGCGATTTCAAGGTGATAGCAGGAGCAAGAGAGA 722  
Db 693 TTGCACTATTGGAATGATATCAACAAGGATTTCCAGCGATTTCCAAAGGGAAAGCAAGTCA 752  
Qy 723 TTTTAAAGCTTGGAGAGTCTTTGTACTATGTGATGTTGTTTACGGCCATATCTGGCA 782  
Db 753 ATATGCAATTGGAGAACTAAGTATATATAGGTACTAGTATGGTGTGCCATTTTGGCA 812  
Qy 783 AGCATTTTGTGGAGCTATTGGGTTGATCTTCTGTGCATCGTCTCTGGTCTCTGGAAT 842  
Db 813 ATTCGCGCTCTGGGGTAGTTGGAGTTATTTTATTTCTCTCTTTACTCTCTGGAAT 872  
Qy 843 TATGGTCAGTCTCTGCTTCCGGTGACGGTATCTTGGCCGTCAATTTGCTTCCAGGAGAA 902  
Db 873 TATAGGCGCTTTTCTTACTTCTGTTACGAAGTTTGGCTGTAATTTTGTTCACGAAAAA 932  
Qy 903 GTTTCAGCGGGGAAAGGTGTCGCTTTGGCTCTCTCCCTCTGGGGATCAGTCTCTTAT 960  
Db 933 --TTTCAGCCGAAAGAGTAGCTATTTTCTAGCTCTATGGGATTTGTTCTATAT 988

Search completed: November 1, 2004, 22:30:03  
Job time : 3748.17 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 1, 2004, 02:40:28 ; Search time 550.844 Seconds  
(without alignments)  
9996.736 Million cell updates/sec

Title: US-09-913-767-2  
Perfect score: 1049  
Sequence: 1 aaagatgaagatgaagacagt.....actgattatgtagcttaaaa 1049

Scoring table: IDENTITY NUC  
Gap 10\_0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_23Sep04:\*  
1: geneseqn1980s:\*  
2: geneseqn1990s:\*  
3: geneseqn2000s:\*  
4: geneseqn2001as:\*  
5: geneseqn2001bs:\*  
6: geneseqn2002as:\*  
7: geneseqn2002bs:\*  
8: geneseqn2003as:\*  
9: geneseqn2003bs:\*  
10: geneseqn2003cs:\*  
11: geneseqn2003ds:\*  
12: geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1049	100.0	1049	3	AA97920 A. thalia
2	528	50.3	1225	3	AA97919 A. thalia
3	428.8	40.9	2586	3	AAC44184 Arabidops
4	199.2	19.0	493	3	AAC36831 Arabidops
5	167.6	16.0	3387	8	AB214407 Arabidops
6	167.6	16.0	3387	8	ADA68246 Arabidops
7	157	15.0	1293	3	AA97922 A. thalia
8	155.8	14.9	1200	8	ADA70627 Rice gene
9	154.2	14.7	1071	3	AA97925 A. thalia
10	149	14.2	1403	3	AAC37652 Arabidops
11	148.8	14.2	1047	6	ADG88190 A. thalia
12	148.8	14.2	1081	3	AA97924 A. thalia
13	142.2	13.6	2660	3	AAC47974 Arabidops
14	137	13.1	1194	3	AA97923 A. thalia
15	133	12.7	83698	6	ABN85767 Arabidops
16	112	10.7	1145	3	AA97921 A. thalia
17	112	10.7	1292	3	AAC51119 Arabidops
18	110.4	10.5	1295	3	AAC37289 Arabidops
19	108.8	10.4	1152	8	ADA70213 Rice gene
20	103.2	9.8	1155	8	ADA69686 Rice gene
21	103	9.8	2175	8	ADA70748 Rice gene

22	90	8.6	1418	3	AAC46395	Aac46395 Arabidops
23	90	8.6	1421	3	AAC39779	Aac39779 Arabidops
24	81.6	7.8	1040	8	ADA69568	Ada69568 Rice gene
25	74.4	7.1	592	10	ABX56692	Abx56692 Arabidops
26	56.6	5.4	1173	12	ADN72520	Adn72520 Thale cre
27	51	4.9	774	6	ABN99127	Abn99127 Arabidops
28	50.2	4.8	477	3	AAC37273	Aac37273 Arabidops
29	48.2	4.6	332	12	ADP92185	Adp92185 Cotton ex
30	42	4.0	677	6	ABQ35095	Abq35095 Oligonucl
31	42	4.0	677	6	ABQ35094	Abq35094 Oligonucl
32	40	3.8	1186	10	ADC87306	Adc87306 Human GPC
33	39.2	3.7	6741	3	AAAI0595	Aai0595 Gene enco
34	39.2	3.7	9447	6	ABL33166	Ab133166 Human inm
35	38	3.6	1648	10	ADC87042	Adc87042 Human GPC
36	38	3.6	2000	8	ADA71938	Ada71938 Rice gene
37	37.8	3.6	18467	4	ABL06976	Ab106976 Drosophil
38	37.4	3.6	2000	8	ADA71938	Ada71938 Rice gene
39	37	3.5	110000	6	ABX08336_05	Continuation (6 of
40	37	3.5	110000	6	ABX08336_06	Continuation (7 of
41	37	3.5	110000	12	ADJ25985_05	Continuation (6 of
42	37	3.5	110000	12	ADJ25985_06	Continuation (7 of
43	37	3.5	110000	12	ADN97989_05	Continuation (6 of
44	37	3.5	110000	12	ADN97989_06	Continuation (7 of
45	37	3.5	110000	12	ADO50281_05	Continuation (6 of

ALIGNMENTS

RESULT 1  
AAA97920  
ID AAA97920 standard; DNA; 1049 BP.  
XX  
AC AAA97920;  
XX  
DT 19-JAN-2001 (first entry)  
XX  
DE A. thaliana PUP1 DNA #2.  
XX  
KW PUP1; transgenic plant; nucleobase transporter; apical dominance;  
KW flowering behaviour; senescence; pesticide distribution; ds.  
XX  
OS Arabidopsis thaliana.  
XX  
PN DE19907209-A1.  
XX  
PD 24-AUG-2000.  
XX  
PF 19-FEB-1999; 99DE-01007209.  
XX  
PR 19-FEB-1999; 99DE-01007209.  
XX  
(FROM/) FROMMER W.  
XX  
Gillissen B, Buerkle L, Andre B, Frommer WB;  
WPI; 2000-566202/53.  
XX  
PT Nucleic acid, useful for producing transgenic plants with altered  
PT nucleobase transport, encodes a nucleobase transporter protein of  
PT Arabidopsis thaliana.  
XX  
PS Claim 2; Page 12; 24pp; German.

CC This invention describes a novel nucleic acid encoding a plant nucleobase  
CC transporter (I). (I) is produced by complementation of a nucleobase  
CC transport (NBT)-defective host cell with a plant gene bank by selection  
CC of NBT-positive cells. (I) is used to isolate homologous sequences from  
CC bacteria, fungi, plants, animals and humans, for expression of the  
CC encoded protein (II) in prokaryotic or eukaryotic cells, for inhibiting  
CC expression of (II) (when in antisense orientation), and to produce  
CC transgenic crop plants. The transgenic plants have modified nucleobase  
CC transport properties, e.g. altered affinity and substrate specificity

CC that may result in more efficient nucleobase transport in leaves, changes  
CC in apical dominance, flowering behaviour and senescence, or improved  
CC distribution of pesticides. This sequence encodes the Arabidopsis  
CC thaliana PUP1 protein which is described in the method of the invention  
xx  
SQ Sequence 1049 BP; 220 A; 239 C; 241 G; 349 T; 0 U; 0 Other;

```
Query Match      100.0%; Score 1049; DB 3; Length 1049;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1049; Conservative 0; Mismatches 0; Indels 0; Gaps 0
```

Qy	1	AAGATGAAGATGAAGACAGTCTCTTGTAATCATAAACTGTATATCTTTGGCCATTGGAAC	60
Db	1	AAGATGAAGATGAAGACAGTCTCTTGTAATCATAAACTGTATATCTTTGGCCATTGGAAC	60
Qy	61	TGTGGAGGCCCTCAATGATGCGTCTCTPACTTCCAAAATGGTGGCGAAAGGATCTCGTTT	120
Db	61	TGTGGAGGCCCTCAATGATGCGTCTCTPACTTCCAAAATGGTGGCGAAAGGATCTCGTTT	120
Qy	121	CCAAGCTTCCTTCAAAACGGTTGGTTGTGCCACTCATTTTTCTTCCTCTCTCTTATCTTTC	180
Db	121	CCAAGCTTCCTTCAAAACGGTTGGTTGTGCCACTCATTTTTCTTCCTCTCTCTTATCTTTC	180
Qy	181	CTCCGCGTCGCTGGTCTTGAAGACAAGAAACGACTCCATTTTCTTCATGAACCT	240
Db	181	CTCCGCGTCGCTGGTCTTGAAGACAAGAAACGACTCCATTTTCTTCATGAACCT	240
Qy	241	CCTCTCTTTATCGCCGCTATCGTTGTTGGTTTGTCTCGTGGGATTTGACAAATTACCTCTAC	300
Db	241	CCTCTCTTTATCGCCGCTATCGTTGTTGGTTTGTCTCGTGGGATTTGNCAAATTACCTCTAC	300
Qy	301	TCTTAGGGTTAGCTTTATATCCCTGTTTCTACTCGGTCCTTTGATCATCTCCGCGCAATT	360
Db	301	TCTTAGGGTTAGCTTTATATCCCTGTTTCTACTCGGTCCTTTGATCATCTCCGCGCAATT	360
Qy	361	GGCTTCACTGCTCTCTTTGCAATTTTTATGTTGTAAGCAAAAGTTTACACTTTTCACTATA	420
Db	361	GGCTTCACTGCTCTCTTTGCAATTTTTATGTTGTAAGCAAAAGTTTACACTTTTCACTATA	420
Qy	421	AACGCTATCGTTTGTCTCACTGGTGGTGGCGTAGTCTCTTGCCCTTAACTCTGATAGTAGC	480
Db	421	AACGCTATCGTTTGTCTCACTGGTGGTGGCGTAGTCTCTTGCCCTTAACTCTGATAGTAGC	480
Qy	481	AAGCTTGAAACGAGACACACAAGGAATATGTTGTTGGGTTCAATCATGACTCTTGTTGCA	540
Db	481	AAGCTTGCAACGAGACACACAAGGAATATGTTGTTGGGTTCAATCATGACTCTTGTTGCA	540
Qy	541	GCTCTTCTATGGGTTTATATGCGACTTGTCCAGCTTTTCTTACAAGAAATCTGGTCTAG	600
Db	541	GCTCTTCTATGGGTTTATATGCGACTTGTCCAGCTTTTCTTACAAGAAATCTGGTCTAG	600
Qy	601	CGAATCAGTATACGCTCCGCTCGAGTTCAGATGGTCTTATGCTTTGCTGCCACTTGT	660
Db	601	CGAATCAGTATACGCTCCGCTCGAGTTCAGATGGTCTTATGCTTTGCTGCCACTTGT	660
Qy	661	GTCTGGCTCGTGGGATGCTAGCCGCTGGCGATTTCAAGGTGATAGCAGAGAGCAAGA	720
Db	661	GTCTGGCTCGTGGGATGCTAGCCGCTGGCGATTTCAAGGTGATAGCAGAGAGCAAGA	720
Qy	721	GATTTTAAAGCTTGGAGAGTCTTTGTACTATGTGTGATTTGTGTTACGGGCCATAAATCTGG	780
Db	721	GATTTTAAAGCTTGGAGAGTCTTTGTACTATGTGTGATTTGTGTTACGGGCCATAAATCTGG	780
Qy	781	CAAGCATTTTTTGTGGGAGCTATTGGGTTGATCTTCTGTGCACTGGTCTCTGGA	840
Db	781	CAAGCATTTTTTGTGGGAGCTATTGGGTTGATCTTCTGTGCACTGGTCTCTGGA	840
Qy	841	ATTATGGTCAGTCTCTGCTCCGGTGACGGTGTATCTTGGCCGTCAATTTGCTTCCAGAG	900
Db	841	ATTATGGTCAGTCTCTGCTCCGGTGACGGTGTATCTTGGCCGTCAATTTGCTTCCAGAG	900
Qy	901	AAGTTTTCAGGCGGGGAAAGGTGTGCTTTGGCTCTCTCCCTCTGGGGATCAGTCTCTTAT	960

Db	901	AAAGTTTCAGCGCGGGAAGAGTGTGCGCTTTGGCTCTCTCCCTCTGGGGATCAGTCTCTTAT	960
Qy	961	TTCTATGGACAGAGTTAAATCCGAGGAGAACTTAAGGCTCAGGATACACAACCTGTCTCAG	1020
Db	961	TTCTATGGACAGAGTTAAATCCGAGGAGAACTTAAGGCTCAGGATACACAACCTGTCTCAG	1020
Qy	1021	CTTCCAGTTACTGATTATGTAGCTTAAAA	1049
Db	1021	CTTCCAGTTACTGATTATGTAGCTTAAAA	1049
RESULT 2			
AAA97919			
ID	AAA97919 standard; DNA; 1225 BP.		
XX			
AC	AAA97919;		
XX			
DT	19-JAN-2001 (first entry)		
XX			
DE	A. thaliana PUP1 DNA #1.		
XX			
KW	PUP1; transgenic plant; nucleobase transporter; apical dominance;		
KW	flowering behaviour; senescence; pesticide distribution; ds.		
XX			
OS	Arabidopsis thaliana.		
XX			
Key	Location/Qualifiers		
PH	25..1095		
CDS	/*tag= a		
FT	/product= "PUP1"		
FT			

Query Match 50.3%; Score 528; DB 3; Length 1295;  
Best Local Similarity 70.6%; Pred. No. 3.2e-157;  
Matches 767; Conservative 0; Mismatches 285; Indels 33; Gaps 3;  
QV 1 AAGATGAAGATGAAGACAGTCTTTGTAATCATAACTGTATATCTTGGCCATTGGAAAC 60

Db 16 AAGAGAGATGAAGATGGTTTGGATAATCATAACTGATTATCTCTCACTATAGGAACA 75  
Qy 61 TGTGGAGGCGCTCTAAATGATGCGTCTCTACTTCCAAAATGGTGGCGAAAGGATCTGCTTT 120  
Db 76 TGTGGAGGCTCTTGTAACTGCTCTACTTCCAAATGGCGGAAACGAATCTGGTTC 135  
Qy 121 CCAAGTCTGCTTCAAAACCGTGTGTGTCACACTCAATTTCTTCCCTCTCTCTTATCTTTC 180  
Db 136 ATGAGCTTCTATCAACCGCGTGTGTTTCCAATCATCTCATCCCTCTCTTGGTCTCTTTC 195  
Qy 181 CTCGCGCGTGTGTTG-----CCTTGAAGAAACAAGACGATCTCA 222  
Db 196 CTCAGCGGTGCGCGCGCAACCGCAACCCCTAACACGCGGAAACAAGCGGAAACAAG 255  
Qy 223 TTTTTCCTCATGAAACCTCTCTCTTTATCGCGCTATCGTTGTTGTTTGTCTGCGGGA 282  
Db 256 CTCCTCTCATGGAACCTCTCTGTTTATCGCTCCATTTGTCTATAGGGTGTCTACAGGA 315  
Qy 283 TTTGACAAATTAACCTTACTCTTACGGGTAGCTTATATCCCTGTTTCTACTCGCTCTTTG 342  
Db 316 CTTGACAACTACTTATATCTTACGGATTAGCATATCTGCCAGTTTCAACTTCATCGCTC 375  
Qy 343 ATCATCTCGCGCAATTAGCTTCACTGCTCTCTTTTGCAATTTTATGTTGAAGCAAAAG 402  
Db 376 ATAATCGGAACCTCAACTAGCTTTCAACGCTCTCTTCGCTTCTTGTGTAGTCAAGCAAAA 435  
Qy 403 TTCACACCTTTCACTATAAAGCGTATCGTTTGTCTCACTGGTGGTCCGTAGTCCCTTGGC 462  
Db 436 TTCACCTCGTCTCCCAATAAAGCGCTGTTTGTGACGGTGTGATACGGGATCCCTTGGC 495  
Qy 463 CTTAACTCTGATAGTCAAGCTTGAACCGAGACACACAAGGAATATGTTGTGGGTTTC 522  
Db 496 TTAACACAGTATGGAGACAAACCGGCTAAGGAGACGAAGAAGATATGTTGGTTGGTTC 555  
Qy 523 ATCATCACTTGGTGCAGCTCTTCTCTATGGTGTATATGTCACCTTGTGACGCTTTC 582  
Db 556 TTGATCACTGTGGTGCAGCTCTTCTCTATGCTTTTATATACCCTGCTGTTGAGCTAACT 615  
Qy 583 TACAAGAAATCTGGTCAGCGAATCAGTATAGCTCGCGCTCGAGTTCAGATGGTCTTA 642  
Db 616 TACAAGAAAGCTGTCAGAAATCACTTTCCACCTTGTGCTTGAGATTCAGATGGTCA 675  
Qy 643 TGCCTTGTGCTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 702  
Db 676 TGCCTTGTGCTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 735  
Qy 703 ATAGCAGGAGACGAAGATTTTAAGCTTGGAGATC---TTTGTACTATGTGGTGATT 759  
Db 736 ATAGCAAGAGAACGAAGAGTTCAAGATTTGGAGGATCAGTGTGTTTACTATGCAATTGATA 795  
Qy 760 GTGTTCAAGCCATAATCTGGCAAGCATTTTGTGGGAGCTATTTGGGTTGATCTCTCTGT 819  
Db 796 GTGATCACAGGAATAATATGGCAAGGTTTCTCTTAGGAGCCATAGGGAATGTGTTTGT 855  
Qy 820 GCATCGTCTCTGCTCTCTGGAATTTAGTTCAGTGTCTGCTTCCGCTGACGGTGATCTTG 879  
Db 856 GCATCATCACTAGCTTCTGCTGTTCTGATAAGTGTCTGCTTCCGCTGACTGAATTTTC 915  
Qy 880 GCGCTCAATTTGCTTCAAGAGAGTTTCAAGCGGGGAAAGGTGTGCTGCTTGTGCTCTCTCC 939  
Db 916 GCGCTGTTTGTGTTCCGGGAGAAGTTTCAAGCAGAGAAAGGTGTCTCTACTCTCTTCT 975  
Qy 940 CTCTGGGATCAGTCTCTTATTTCTATGACAGGTTAATCC-----GAGGAG 987  
Db 976 CTTTGGGATTTGTCTCTTACTTCTACGGCGAGTTTAAATCCGGCAAGAAAGTTGTGAT 1035  
Qy 988 AAGACTAAGCGCTCAGGATACAACTGTCTCAGCTTCCAGTTTACTGATTTAGTCTTAA 1047  
Db 1036 AAACCTCAACCGCGGAGACAGACTGCTTATTTCTTCCAGTTAGTATATGTTGCTTAA 1095

AAC4184  
ID AAC4184 standard; DNA; 2586 BP.  
XX  
AC AAC4184;  
XX  
DT 18-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 41938.  
XX  
KW Hybridisation assay; genetic mapping; gene expression control;  
KW protein identification; signal transduction pathway; metabolic pathway;  
KW promoter; termination sequence; ss.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-00301439.  
XX  
PR 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.

PR 18-JUN-1999;	99US-0139460P.	PR 23-AUG-1999;	99US-0149902P.
PR 18-JUN-1999;	99US-0139461P.	PR 23-AUG-1999;	99US-0149930P.
PR 18-JUN-1999;	99US-0139462P.	PR 26-AUG-1999;	99US-0150566P.
PR 18-JUN-1999;	99US-0139463P.	PR 26-AUG-1999;	99US-0150884P.
PR 18-JUN-1999;	99US-0139750P.	PR 27-AUG-1999;	99US-0151065P.
PR 18-JUN-1999;	99US-0139763P.	PR 27-AUG-1999;	99US-0151066P.
PR 21-JUN-1999;	99US-0139817P.	PR 27-AUG-1999;	99US-0151080P.
PR 22-JUN-1999;	99US-0139899P.	PR 30-AUG-1999;	99US-0151303P.
PR 23-JUN-1999;	99US-0140353P.	PR 31-AUG-1999;	99US-0151438P.
PR 23-JUN-1999;	99US-0140354P.	PR 01-SEP-1999;	99US-0151930P.
PR 24-JUN-1999;	99US-0140695P.	PR 07-SEP-1999;	99US-0152363P.
PR 28-JUN-1999;	99US-0140823P.	PR 10-SEP-1999;	99US-0153070P.
PR 29-JUN-1999;	99US-0140991P.	PR 13-SEP-1999;	99US-0153758P.
PR 30-JUN-1999;	99US-0141287P.	PR 15-SEP-1999;	99US-0154018P.
PR 01-JUL-1999;	99US-0141842P.	PR 16-SEP-1999;	99US-0154039P.
PR 01-JUL-1999;	99US-0142154P.	PR 20-SEP-1999;	99US-0154779P.
PR 02-JUL-1999;	99US-0142055P.	PR 22-SEP-1999;	99US-0155139P.
PR 06-JUL-1999;	99US-0142390P.	PR 23-SEP-1999;	99US-0155486P.
PR 08-JUL-1999;	99US-0142803P.	PR 24-SEP-1999;	99US-0155659P.
PR 09-JUL-1999;	99US-0142920P.	PR 28-SEP-1999;	99US-0156458P.
PR 12-JUL-1999;	99US-0142977P.	PR 29-SEP-1999;	99US-0156596P.
PR 13-JUL-1999;	99US-0143542P.	PR 04-OCT-1999;	99US-0157117P.
PR 14-JUL-1999;	99US-0143624P.	PR 05-OCT-1999;	99US-0157753P.
PR 15-JUL-1999;	99US-0144005P.	PR 06-OCT-1999;	99US-0157865P.
PR 16-JUL-1999;	99US-0144085P.	PR 07-OCT-1999;	99US-0158029P.
PR 16-JUL-1999;	99US-0144086P.	PR 08-OCT-1999;	99US-0158232P.
PR 19-JUL-1999;	99US-0144325P.	PR 12-OCT-1999;	99US-0158369P.
PR 19-JUL-1999;	99US-0144331P.	PR 13-OCT-1999;	99US-0159293P.
PR 19-JUL-1999;	99US-0144332P.	PR 13-OCT-1999;	99US-0159294P.
PR 19-JUL-1999;	99US-0144333P.	PR 13-OCT-1999;	99US-0159295P.
PR 19-JUL-1999;	99US-0144334P.	PR 14-OCT-1999;	99US-0159329P.
PR 19-JUL-1999;	99US-0144335P.	PR 14-OCT-1999;	99US-0159330P.
PR 20-JUL-1999;	99US-0144352P.	PR 14-OCT-1999;	99US-0159331P.
PR 20-JUL-1999;	99US-0144632P.	PR 14-OCT-1999;	99US-0159637P.
PR 20-JUL-1999;	99US-0144884P.	PR 14-OCT-1999;	99US-0159638P.
PR 21-JUL-1999;	99US-0144814P.	PR 18-OCT-1999;	99US-0159584P.
PR 21-JUL-1999;	99US-0145086P.	PR 21-OCT-1999;	99US-0160741P.
PR 21-JUL-1999;	99US-0145088P.	PR 21-OCT-1999;	99US-0160767P.
PR 22-JUL-1999;	99US-0145089P.	PR 21-OCT-1999;	99US-0160768P.
PR 22-JUL-1999;	99US-0145087P.	PR 21-OCT-1999;	99US-0160770P.
PR 22-JUL-1999;	99US-0145089P.	PR 21-OCT-1999;	99US-0160814P.
PR 22-JUL-1999;	99US-0145192P.	PR 21-OCT-1999;	99US-0160815P.
PR 23-JUL-1999;	99US-0145145P.	PR 22-OCT-1999;	99US-0160980P.
PR 23-JUL-1999;	99US-0145218P.	PR 22-OCT-1999;	99US-0160981P.
PR 23-JUL-1999;	99US-0145224P.	PR 22-OCT-1999;	99US-0160989P.
PR 26-JUL-1999;	99US-0145276P.	PR 25-OCT-1999;	99US-0161404P.
PR 27-JUL-1999;	99US-0145913P.	PR 25-OCT-1999;	99US-0161405P.
PR 27-JUL-1999;	99US-0145918P.	PR 25-OCT-1999;	99US-0161406P.
PR 27-JUL-1999;	99US-0145919P.	PR 26-OCT-1999;	99US-0161359P.
PR 28-JUL-1999;	99US-0145951P.	PR 26-OCT-1999;	99US-0161360P.
PR 02-AUG-1999;	99US-0146386P.	PR 26-OCT-1999;	99US-0161361P.
PR 02-AUG-1999;	99US-0146388P.	PR 28-OCT-1999;	99US-0161920P.
PR 02-AUG-1999;	99US-0146389P.	PR 28-OCT-1999;	99US-0161921P.
PR 03-AUG-1999;	99US-0147038P.	PR 28-OCT-1999;	99US-0161992P.
PR 04-AUG-1999;	99US-0147204P.	PR 29-OCT-1999;	99US-0161993P.
PR 04-AUG-1999;	99US-0147302P.	PR 29-OCT-1999;	99US-0162142P.
PR 05-AUG-1999;	99US-0147192P.		
PR 05-AUG-1999;	99US-0147260P.		
PR 06-AUG-1999;	99US-0147303P.		
PR 06-AUG-1999;	99US-0147416P.		
PR 09-AUG-1999;	99US-0147493P.		
PR 09-AUG-1999;	99US-0147935P.		
PR 10-AUG-1999;	99US-0148171P.		
PR 11-AUG-1999;	99US-0148319P.		
PR 12-AUG-1999;	99US-0148341P.		
PR 13-AUG-1999;	99US-0148565P.		
PR 13-AUG-1999;	99US-0148684P.		
PR 16-AUG-1999;	99US-0149368P.		
PR 17-AUG-1999;	99US-0149175P.		
PR 18-AUG-1999;	99US-0149426P.		
PR 20-AUG-1999;	99US-0149722P.		
PR 20-AUG-1999;	99US-0149723P.		
PR 20-AUG-1999;	99US-0149929P.		

Query Match

Best Local Similarity

Matches 589; Conservative

Score 428.8; DB 3; Length 2586;

Pred. No. 2.2e-125;

Mismatches 209; Indels 15; Gaps 2;

Qy	249	TATCGCGCTATCGTTCCTGGTTTGGTTCGTCGTTGGGATTTGACAAATTAACCTTACTCTTACGG	308
Db	1646	TATCGCTCCATTGTCATAGGTTGCTCAGAGACTTGACAACTACTTATATCTTTACGG	1705
Qy	309	GTTAGCTTATATCCCTCTTCTTACTGCGTCTTTTGATCATCTCCGCGCAATTAGCTTCAAC	368
Db	1706	ATTAGCATATCTGCAGCTTTCACCTTCATCGTCATAATCGGAACCTCAACTAGCTTTCAA	1765
Qy	369	TGCTCTCTTTGCAATTTTATGGTGAAGCAAAAGTTACACCTTTTCACTATAAAGCTAT	428
Db	1766	CGCTCTCTTCGCTTCTTCTTTAGTCAAGCAAAAGTTCACTCCGTTCTCCATAAAGCGGT	1825
Qy	429	CGTTTGTCTCACTGGTGGTGGCGGTAGTCTTCCTTAACTCTGATAGTGACAAGCTTGC	488



Db	1826	CGTTTGTGACGGTTGGTATCGGGATCCTTGGCTTACACAGTGATGGAGACAAACCGGC	1885	PR	06-APR-1999;	99US-0128234P.
QY	489	AAACGAGACACAAAGGAATATGTTGTGGTTTCATCATGACTCTTGGTGACGCTCTTCT	548	PR	08-APR-1999;	99US-0128714P.
Db	1886	TAAGGAGACGAGAAGATGATGGTTGGTTCTTGAAGACTGTGGTGACGCTCTTCT	1945	PR	16-APR-1999;	99US-0129845P.
QY	549	CTATGGGTTTATATGCCACTTTCGAGCTTCTTACAAGAAATCTGGTCAGCGAATCAC	608	PR	19-APR-1999;	99US-0130077P.
Db	1946	CTATGCTTTTATATACCGCTCGTTGAGCTAACTTACAAGAAAGCTCGTCAAGAAATCAC	2005	PR	21-APR-1999;	99US-0130443P.
QY	609	GTATAGCTCGCCTCGAGTTCAGATGTTCTATGCTTGTGTCGACCTTGTCTGCTCT	668	PR	23-APR-1999;	99US-0130510P.
Db	2006	TTTCCCACTTGTGCTTGAATTCAGATGTTGATGCTGCTGCTGCTGCTGCTGCTGCT	2065	PR	28-APR-1999;	99US-0130891P.
QY	669	CGTGGGATGCTAGCGCTGCGCATTTCAAGTGATGACGAGAGCAAGAGATTTTAA	728	PR	30-APR-1999;	99US-0131449P.
Db	2066	CATTGGCATGTTTCATCGTTGGAGATTTAAGGTGATGACGAGAGCAAGAGATTTCAA	2125	PR	99US-0132407P.	99US-0132484P.
QY	729	GCTTGGAGATC---TTTGTACTATGTGGTGAATTGTTTCAACGCCATATCTGGCAAGC	785	PR	04-MAY-1999;	99US-0132485P.
Db	2126	GATTGGAGGATCAGTGTTTTACTATGATGATGATCAGAGGAATATATGGCAAGG	2185	PR	05-MAY-1999;	99US-0132485P.
QY	786	ATTTTGTGGAGCTATTTGGGTTGATCTTCTGTGATCGTCTGCTGCTGCTGCTGCTGCT	845	PR	06-MAY-1999;	99US-0132486P.
Db	2186	TTTCTTCTTAGGAGCCATAGGATTTGTTTGTGATCATCACTAGCTTCTGCTGCTCT	2245	PR	07-MAY-1999;	99US-0132487P.
QY	846	GCTCAGTCTCTGCTCCGCTGACGCTGATCTTGGCGCTCATTTGCTTCCAGGAGATTT	905	PR	11-MAY-1999;	99US-013256P.
Db	2246	GATAAGTGTCTGCTTCCGCTGACTGAAGTTTTCGCGCTGCTGCTGCTGCTGCTGCT	2305	PR	14-MAY-1999;	99US-0134218P.
QY	906	TCAGGCGGGGAAAGTGTGCTTGGCTCTCTCCCTCTGGGGATCAGTCTCTTATTCTTA	965	PR	14-MAY-1999;	99US-0134221P.
Db	2306	TCAGGCGAGAAAGTGTCTCTACTCTTCTTCTTGGGATTTGCTCTTACTTCTTA	2365	PR	14-MAY-1999;	99US-0134370P.
QY	966	TGACACAGGTTAAATCC-----GAGGAGAAAGCTAAGGCTCAGGATACAACT	1013	PR	18-MAY-1999;	99US-0134768P.
Db	2366	CGGCGAGTTTAAATCCGGAAGAAAGTTGTGATAAACCTCAACCGCGAGACAGAACT	2425	PR	19-MAY-1999;	99US-0134941P.
QY	1014	GTCTCAGCTTCCAGTTACTGATTATGACTTAA 1047		PR	20-MAY-1999;	99US-0135124P.
Db	2426	GCCTATTTCTCCAGTTAGTATTGTTGCTTAA 2459		PR	21-MAY-1999;	99US-0135353P.
RESULT 4						99US-0136392P.
AAC36831						99US-0136782P.
ID	AAC36831 standard; DNA; 493 BP.					99US-0137222P.
XX	AAC36831;					99US-0137528P.
XX	17-OCT-2000 (first entry)					99US-0137724P.
DT	Arabidopsis thaliana DNA fragment SEQ ID NO: 15222.					99US-0138094P.
DE	Hybridisation assay; genetic mapping; gene expression control;					99US-0138540P.
KW	protein identification; signal transduction pathway; metabolic pathway;					99US-0138847P.
KW	promoter; termination sequence; ss.					99US-0139119P.
OS	Arabidopsis thaliana.					99US-0139453P.
XX	EP1033405-A2.					99US-0139454P.
PN	06-SEP-2000.					99US-0139455P.
PD	25-FEB-2000; 2000BP-00301439.					99US-0139457P.
XX	25-FEB-1999; 99US-0121825P.					99US-0139458P.
PR	05-MAR-1999; 99US-0123180P.					99US-0139459P.
PR	09-MAR-1999; 99US-0123548P.					99US-0139460P.
PR	23-MAR-1999; 99US-0125788P.					99US-0139461P.
PR	25-MAR-1999; 99US-0126264P.					99US-0139462P.
PR	29-MAR-1999; 99US-0126785P.					99US-0139463P.
PR	01-APR-1999; 99US-0127462P.					99US-0139463P.



DE Arabidopsis thaliana stress regulated gene SEQ ID NO 2212.  
XX Arabidopsis thaliana; plant; gene; stress; transgenic; ds.  
KW Arabidopsis thaliana.  
OS Arabidopsis thaliana.  
XX WO200216655-A2.  
PN 28-FEB-2002.  
XX 24-AUG-2001; 2001WO-US026685.  
XX 24-AUG-2000; 2000US-0227866P.  
PR 26-JAN-2001; 2001US-0264647P.  
PR 22-JUN-2001; 2001US-0300111P.  
XX (SCRI ) SCRIPPS RES INST.  
PA (SYGN ) SYNGENTA PARTICIPATIONS AG.  
PA Harper JF, Kreps J, Wang X, Zhu T;  
XX WPI; 2002-304127/34.  
XX Identifying a stress condition to which a plant cell has been exposed and  
PT producing plants with increased tolerance to these abiotic stresses.  
XX Claim 144; SEQ ID NO 2212; 577bp + Sequence Listing; English.  
XX The invention relates to identifying a stress condition to which a plant  
CC cell has been exposed, comprising: (a) contacting nucleic acid  
CC representative of expressed polynucleotides in the plant cell with an  
CC array or probes representative of the plant cell genome; and (b)  
CC detecting a profile of expressed polynucleotides in the plant cell  
CC characteristic of a stress response. The method is useful in the  
CC production of transgenic plants, cells and seeds and in producing plants  
CC with increased tolerance to abiotic stress. The present sequence is that  
CC of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used  
CC in methods of the invention. Note: The sequence data for this patent is  
CC not represented in the printed specification but is based on sequence  
CC information supplied to Derwent by the European Patent Office  
XX Sequence 3387 BP; 870 A; 781 C; 654 G; 1082 T; 0 U; 0 Other;

Query Match 16.0%; Score 167.6; DB 6; Length 3387;  
Best Local Similarity 48.6%; Pred. No. 5.7e-42;  
Matches 492; Conservative 0; Mismatches 514; Indels 6; Gaps 1;  
QY 1 AAGATGAAGATGAGACAGTCTTGTGAATCATAACTGTATATCTTGGCCCATTTGAAC 60  
DB 2353 AAGAAGCTGTAAGAGGTGGCTCGGTCTCCATATACGCAATCTTGTCTATCTTCTGCCAA 2412  
QY 61 TGTGAGGCCCTCTAATGATGGTCTCTACTTCCAAATGGTGGCAAGGATCTGGTTT 120  
DB 2413 CCACCTGTACAGTCTCGGTGAGACTGTACTATGAAATGGAGGAAAGCAATATG 2472  
QY 121 CCAAGCTCTCTCAAAACCGTTGGTTGTCCCACTATTTCTTCCCTCTCTTATCTTTC 180  
DB 2473 GTAACACTTCTTCAACTCATTTGGCTCCCTGTACTGATTTCTCTCGCTTCTTCTCGA 2532  
QY 181 CTCGCGCGTGTGGTTCCTGAAGAACAAAGAACGACTCCATTTTTCCTCATGAACCT 240  
DB 2533 ATCAGGCAACCCAAATCAACAGATACAAATTTTCAGTCAGTCCCTTCTCCACCCCTT 2592  
QY 241 CTTCTCTTTATGCCGCTATCTGTTGGTTTGTCTGGGATTTGACAAATTAACCTTAC 300  
DB 2593 GCATCGGTTTAC-----TGTGCACTGGACTGTAGTCTGTCTATGCTTATTGTCT 2646  
QY 301 TCTTAGCGGTAGCTTATATCCCTGTTTCTACTGCGTCTTTGATCATCTCCGCGCAATTA 360  
DB 2647 GCAGTTGGGTGCTTTACTTACCAAGTCTCTACTTTCTCCCTCATCTTGGCCCTCACAGTTG 2706  
QY 361 GGCCTTCACTGCTCTCTTTGCATTTTATGGTGAAGCAAAAGTTCACACCTTTTCACTATA 420

DB 2707 GCCTTCACTGCCTTTTCTCATATATTTCTTAACGCAAAAGTTCACTCTCTTTGTAGTGC 2766  
QY 421 AAGCGTATCGTTTTCCTCACTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 480  
DB 2767 AATTCTTTGTTTCTCTTACGGTTTCTCTGCGCTTCTCTGCTGGTCAACACTGATTGAGAA 2826  
QY 481 AAGCTTGGCAACGAGACACACAAGGAATATGTTGTTGGTTCATCATGACTCTTGGTGA 540  
DB 2827 AACACACAAATGTTATCTAGAGTACAATATGTTGGTTCATCTGTACCACTTGGTGT 2886  
QY 541 GCTCTTCTATGGGTTTATATTTGACACTTGTGAGACTTCTTACAGAAATCTGGTCA 600  
DB 2887 TCCGCTGGGATTTGGTATCTATCTCTGATACAACTGCTCTTCCAGGAAGTTTTCAG 2946  
QY 601 CGAATCAGTATACGCTGCGCTCGAGTTCAGATGGTCTTATGCTTTGCTGCTCCACTTCT 660  
DB 2947 AAGCATACATCTCTGAGCAGTCTGCACTTGGCAATTAACAGTCTCTAGTTGCAACTTGT 3006  
QY 661 GTCTGCTGCTGGGGATGCTAGCCGCTGGCGATTTCAAGGTGATAGCAGGAGAAACA 720  
DB 3007 GTGGTACTCATAGGACTGTTTGCAGGTGGAGAGTCTGCAAGTGCAGATGAGA 3066  
QY 721 GATTTAAGCTTGGAGAGTCTTTGTACTATGTTGGTGTATGTTTCAAGGCAATATCTGG 780  
DB 3067 AACTACAACTGGGGAAGTGTCTATATCTTGACTTTGGCCTCAGCAGCTATTTTCTGG 3126  
QY 781 CAAGCATTTTGTGGAGCTATTGGGTTGATCTTCTGTGCATCTCTCTGCTCTCTGGA 840  
DB 3127 CAAGTATACACTGTTGGTGTGGGATTAATCTTCGAGTCTTCTTGTGTCTTCTCAAT 3186  
QY 841 ATTATGCTCAGTCTCTGCTTCCGCTGACGGTGTATCTTGGCGCTCATTTTGTCTCCAGGAG 900  
DB 3187 TCCATAACTGCTGTGGGACTGCTATAGTTCAGTGTGAGCAGTATAGTTTTCATGAT 3246  
QY 901 AAGTTTACGGGGGAAAGGTGTCCTTGGCTCTCTCCCTCTGGGATCAGTCTCTTAT 960  
DB 3247 AAGATGGATGATCAAGATTTTCTCCATCATTTTGTAGCTATCTGGGGCTTCTCTTCA 3306  
QY 961 TTCTATGACAGGTTAAATCCGAGGAGAAAGTAAAGGCTCAGGATACACAAC 1012  
DB 3307 GTCTATCAGCTACTCTCGACGAAAGGTTGAAGACTTGCACAGCAAAAC 3358  
RESULT 6  
ADA68246  
ID ADA68246 standard; DNA; 3387 BP.  
XX  
AC ADA68246;  
XX  
DT 20-NOV-2003 (first entry)  
XX  
DE Arabidopsis thaliana gene, SEQ ID 908.  
XX  
KW Plant; bacterial infection; fungal infection; viral infection; ds.  
XX Arabidopsis thaliana.  
XX  
PN WO2003000898-A1.  
XX  
PD 03-JAN-2003.  
XX  
PF 22-JUN-2001; 2001WO-IB001105.  
XX  
PR 22-JUN-2001; 2001WO-IB001105.  
XX  
PA (SYGN ) SYNGENTA PARTICIPATIONS AG.  
XX  
PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;  
PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;  
XX  
DR WPI; 2003-175290/17.  
XX  
PT Identifying at least one gene involved in plant resistance or response to





Db 667 ATCCCGCTCGTCAGGTCGCCACGCGGAGGTACGGGCGCGCAGCGGGCCCGCGGCGAGG 726  
Qy 606 -----CAGGTATACGCTCGCGCTCGAGTTCAGATGGTCTTATGCTTTGCTGCCACT 657  
Db 727 GTGCGCGCTCCCTACGCGCAGCGTGATGCAGATGCAGGCGGTGATGGGCGCGGCGGACG 786  
Qy 658 TGTGTCGCTCGTGGGATGCTAGCCGCTGGCGATTTCAAGGTGATAGCAGGAGAACA 717  
Db 787 GCGGTGTCGTGCTCGCATCGCATCAAGGGCGGACTTCCAGGCGGTGGCGCGGAAGCG 846  
Qy 718 AGAGATTTTAAGCTTGGAGAGTCTTTGTACTATGTGGTGATTTGTTCAAGCGCATATC 777  
Db 847 GCGGGTTGCGGCTCGCGCGCGCCNACTACTACCTCGTCTCGCTGGGACGCGGTGTCG 906  
Qy 778 TGGCAAGCATTTTGTGGGAGCTATTGGGTGATCTTCTGTGTCATCGTCTCGTCTCT 837  
Db 907 TGGCAGCTGCTCAACCTGGGCATCATGGGCTCATCACCTGCGGTGCTGCTGCTGCC 966  
Qy 838 GGAATTATGGTCAAGTCTGCTGCTCCGCTGAGGTGATCTTGGCGGTCAATTTGCTTCAG 897  
Db 967 GGCATCATGATCGCGGTGCTCTCGCGCTCTCGCAGGTCTCTCGCGGTCAATCTCTCCAC 1026  
Qy 898 GAGAAGTTTCAGCGGGGAAAGGTGTCGCTTTGGCTCTCTCCCTCTGGGGATCAGTCTCT 957  
Db 1027 GAGAAGTTTCAGCGGACGAGGACATCGGCTGCTGCTCTGCTCTGGGGATTCGCTCC 1086  
Qy 958 TATTCTTATGACAGTTTAAATCCGAGGAGAAG 990  
Db 1087 TACCTCTACGCGGAGAGCGCGAGAAGAAG 1119

RESULT 9

AAA97925  
ID AAA97925 standard; DNA; 1071 BP.  
AC AAA97925;  
XX  
DT 19-JAN-2001 (first entry)  
DE A. thaliana PUP1 DNA #7.  
XX  
KW PUP1; transgenic plant; nucleobase transporter; apical dominance;  
KW flowering behaviour; senescence; pesticide distribution; ds.  
XX Arabidopsis thaliana.  
OS  
XX DE19907209-A1.  
XX  
XX 24-AUG-2000.  
XX  
XX 19-FEB-1999; 99DE-01007209.  
XX  
XX 19-FEB-1999; 99DE-01007209.  
XX (FROM/) FROMMER W.  
XX  
XX Gillissen B, Buerkle L, Andre B, Frommer WB;  
XX WPI; 2000-566202/53.  
XX  
XX Nucleic acid, useful for producing transgenic plants with altered  
PT nucleobase transport, encodes a nucleobase transporter protein of  
PT Arabidopsis thaliana.  
XX  
PS Claim 2; Page 15; 24pp; German.

CC This invention describes a novel nucleic acid encoding a plant nucleobase  
CC transporter (I). (I) is produced by complementation of a nucleobase  
CC transport (NBT)-defective host cell with a plant gene bank by selection  
CC of NBT-positive cells. (I) is used to isolate homologous sequences from  
CC bacteria, fungi, plants, animals and humans, for expression of the  
CC encoded protein (II) in prokaryotic or eukaryotic cells, for inhibiting  
CC expression of (II) (when in antisense orientation), and to produce

CC transgenic crop plants. The transgenic plants have modified nucleobase  
CC transport properties, e.g. altered affinity and substrate specificity  
CC that may result in more efficient nucleobase transport in leaves, changes  
CC in apical dominance, flowering behaviour and senescence, or improved  
CC distribution of pesticides. This sequence encodes the Arabidopsis  
CC thaliana PUP1 protein which is described in the method of the invention  
XX  
SQ Sequence 1071 BP; 307 A; 236 C; 208 G; 320 T; 0 U; 0 Other;

Query Match 14.7%; Score 154.2; DB 3; Length 1071;  
Best Local Similarity 49.6%; Pred. No. 5.6e-38;  
Matches 458; Conservative 0; Mismatches 453; Indels 12; Gaps 2;  
Qy 49 GCCATTGGAACTGCTGGAGGCCCTCTAAATGATGGCTCTACTTCCAAAATGGTGGCGAA 108  
Db 46 GTCAATGGCCAAATCAGTTGCTACAAATTCGGGCAGACTATACTATGAAAATGGAGGAAC 105  
Qy 109 AGGATCTGGTTCCTCAAGCTTCCAAACCGTTGGTTGTCCACTCATTTTCTTCCCTCTT 168  
Db 106 AGCAATGGCTTAGCAACGGTAGTTACGCTTGTAGGCTTTCCTATTTCTACTTCCATATCAT 165  
Qy 169 CTCTTATCTTTCCTCCCGCTCGTTCGCTTGAAGAACAAAGAAACGACTCCCATTTTTC 228  
Db 166 CTCTTGTCTCAAAACACAT-----ACAACTACAGAGATGCAAAATTAAC 215  
Qy 229 CTGATGAAACCTCTCTCTTTATCGCGCTATCGTTGGTTGCTCGTGGGATTTGAC 288  
Db 216 CTCACTTAGGAACCGTGCAITTA--GTTTACATATGCTTGGACTTCTTGTAGGACGCT 273  
Qy 289 AATTACCTCTACTTTACGGGTTAGCTTATATCCCTGTTTCTACTGGCTCTTTGATCATC 348  
Db 274 TGTACCTATATTCCATTGGACTGCTTTACCTACTGTTTCTACCTTTCCCTGATCTGT 333  
Qy 349 TCGCGCAATTAGGCTTCACTGCTCTCTTTGCAATTTTATGTTGAAGCAAAAGTTCCACA 408  
Db 334 GCATCACAGTTAGCTTTCACCGCTTCTCTCTTATTTACTCAACTCACAATAAACTTACT 393  
Qy 409 CTTTCTACTATAACGCTATCGTTTCTCAGCTGGTGGTGGCTGCTTGCCTTCCCTTAAC 468  
Db 394 CCTATCATTTTGAATTTCTTTTCCCTCACTATATCTTCCACCTCTTGTGCAATTAAT 453  
Qy 469 TCTGATAGTCACAAGCTTGCAAAAGCAGACACACAAGGAATATGTTGTGGTTTCATCATG 528  
Db 454 AACGAGGAATCAGATTTCCAAAAGTTTCAAAAGGAGATATGTCAAAAGTTTTCGATGC 513  
Qy 529 ACTCTTGGTGCAGCTCTTCTATGGGTTTATATGCCACTTTGTGAGCTTTTCTTACAAG 588  
Db 514 ACCGTTGGTGATCTGCTGGGTTTGGTCTACTCTTATCCCTACAAACAGCTAGCCTTTT 573  
Qy 589 AATCTGCTCAGGAATCAGTATACGCTCGCGCTCGAGTTCAGATGGTCTTATGCTTT 648  
Db 574 AAAAGTTTAAAGAAAGCAAACTTTCTCAGAAAGTTATAAATATGATTAATCTACATGAGTCTA 633  
Qy 649 GCTGCCACTTGTGCTCGCTCGGGGATGCTAGCCGCTGCGGATTTCAAGGTGATAGCA 708  
Db 634 GTGCCAGTTGTGTAGCTGGTGGGCTTTTGTCTAGTAGCGAGTGGAATAAATTTGAGC 693  
Qy 709 GGAGAAGCAAGAGATTTTAAAGCTTGGAGAGTCTTTGTACTATATGTTGGTGTGTTCAAG 768  
Db 694 AGTGAATGGAAAACACAAAACCTTGGGAAGTATCTTATGTCATGAACCTAGTGTGAGCA 753  
Qy 769 GCCATATCTGGCAGCAATTTTGTGGGAGCTATTTGGGTTGATCTTCTGTGCATCTGCT 828  
Db 754 GCTGTTACCTGGCAGGATTTTCTCCATCGGTTGACAGGACTGATCTTTCGAGCTTTCCTCC 813  
Qy 829 CTGCTCTCTGGAATTAATGCTCAGTCTCTGCTTCGGGTGACGGTATCTTGGCGCTCAT 888  
Db 814 CTATTTCAAATGCAATAGCGCTTTGGGACTCCCGCTGGTTCCTATCTCGCTGCTCATC 873  
Qy 889 TGCTTCCAGGAGAGTTTTCAGGCGGGGAAAGGTGTGCTGCTTGGCTCTCTCTCCCTCGGGGA 948  
Db 874 ATTTTCCATGACAAAATGAACGGCTTAAAGGTGATTTCTATGATTTCTAGCTATTTGGGGT 933

Qy 949 TCAGTCTCTATTTCATGACA 971  
Db 934 TTCGTATCCTATGCTACCAACA 956

## RESULT 10

AAC37652

ID AAC37652 standard; DNA; 1403 BP.

XX

AAC37652;

XX

17-OCT-2000 (first entry)

XX

DE Arabidopsis thaliana DNA fragment SEQ ID NO: 18172.

XX

KW Hybridisation assay; Genetic mapping; gene expression control;

KW protein identification; signal transduction pathway; metabolic pathway;

KW promoter; termination sequence; ss.

XX

OS Arabidopsis thaliana.

XX

EP1033405-A2.

XX

PD 06-SEP-2000.

XX

PF 25-FEB-2000; 2000EP-00301439.

XX

PR 25-FEB-1999; 99US-0121825P.

PR

PR 05-MAR-1999; 99US-0123180P.

PR

PR 09-MAR-1999; 99US-0123548P.

PR

PR 23-MAR-1999; 99US-0125788P.

PR

PR 25-MAR-1999; 99US-0126264P.

PR

PR 29-MAR-1999; 99US-0126785P.

PR

PR 01-APR-1999; 99US-0127462P.

PR

PR 06-APR-1999; 99US-0128234P.

PR

PR 08-APR-1999; 99US-0128714P.

PR

PR 16-APR-1999; 99US-0129845P.

PR

PR 19-APR-1999; 99US-0130077P.

PR

PR 21-APR-1999; 99US-0130449P.

PR

PR 23-APR-1999; 99US-0130510P.

PR

PR 23-APR-1999; 99US-0130891P.

PR

PR 28-APR-1999; 99US-0131449P.

PR

PR 30-APR-1999; 99US-0132048P.

PR

PR 30-APR-1999; 99US-0132407P.

PR

PR 04-MAY-1999; 99US-0132484P.

PR

PR 05-MAY-1999; 99US-0132485P.

PR

PR 06-MAY-1999; 99US-0132486P.

PR

PR 06-MAY-1999; 99US-0132487P.

PR

PR 07-MAY-1999; 99US-0132863P.

PR

PR 11-MAY-1999; 99US-0134256P.

PR

PR 14-MAY-1999; 99US-0134218P.

PR

PR 14-MAY-1999; 99US-0134219P.

PR

PR 14-MAY-1999; 99US-0134221P.

PR

PR 14-MAY-1999; 99US-0134370P.

PR

PR 18-MAY-1999; 99US-0134768P.

PR

PR 19-MAY-1999; 99US-0134941P.

PR

PR 20-MAY-1999; 99US-0135124P.

PR

PR 21-MAY-1999; 99US-0135353P.

PR

PR 24-MAY-1999; 99US-0135629P.

PR

PR 25-MAY-1999; 99US-0136021P.

PR

PR 27-MAY-1999; 99US-0136392P.

PR

PR 28-MAY-1999; 99US-0136782P.

PR

PR 01-JUN-1999; 99US-0137222P.

PR

PR 03-JUN-1999; 99US-0137528P.

PR

PR 04-JUN-1999; 99US-0137502P.

PR

PR 07-JUN-1999; 99US-0137724P.

PR

PR 08-JUN-1999; 99US-0138094P.

PR

PR 10-JUN-1999; 99US-0138540P.

PR

PR 10-JUN-1999; 99US-0138847P.

PR

PR 14-JUN-1999; 99US-0139119P.

PR

PR 16-JUN-1999; 99US-0139452P.

PR

PR 16-JUN-1999; 99US-0139453P.

PR

PR 17-JUN-1999; 99US-0139492P.

PR

PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.

PR	16-AUG-1999;	99US-0149368P.	
PR	17-AUG-1999;	99US-0149175P.	
PR	18-AUG-1999;	99US-0149426P.	
PR	20-AUG-1999;	99US-0149722P.	
PR	20-AUG-1999;	99US-0149723P.	
PR	20-AUG-1999;	99US-0149929P.	
PR	23-AUG-1999;	99US-0149902P.	
PR	23-AUG-1999;	99US-0149930P.	
PR	25-AUG-1999;	99US-0150566P.	
PR	26-AUG-1999;	99US-0150884P.	
PR	27-AUG-1999;	99US-0151065P.	
PR	27-AUG-1999;	99US-0151066P.	
PR	27-AUG-1999;	99US-0151080P.	
PR	30-AUG-1999;	99US-0151303P.	
PR	31-AUG-1999;	99US-0151438P.	
PR	01-SEP-1999;	99US-0151930P.	
PR	07-SEP-1999;	99US-0152363P.	
PR	10-SEP-1999;	99US-0153070P.	
PR	13-SEP-1999;	99US-0153758P.	
PR	15-SEP-1999;	99US-0154018P.	
PR	16-SEP-1999;	99US-0154039P.	
PR	20-SEP-1999;	99US-0154779P.	
PR	22-SEP-1999;	99US-0155139P.	
PR	23-SEP-1999;	99US-0155486P.	
PR	24-SEP-1999;	99US-0155659P.	
PR	28-SEP-1999;	99US-0156458P.	
PR	29-SEP-1999;	99US-0156596P.	
PR	04-OCT-1999;	99US-0157117P.	
PR	05-OCT-1999;	99US-0157533P.	
PR	06-OCT-1999;	99US-0157865P.	
PR	07-OCT-1999;	99US-0158029P.	
PR	08-OCT-1999;	99US-0158232P.	
PR	12-OCT-1999;	99US-0158369P.	
PR	13-OCT-1999;	99US-0159293P.	
PR	13-OCT-1999;	99US-0159294P.	
PR	13-OCT-1999;	99US-0159295P.	
PR	14-OCT-1999;	99US-0159329P.	
PR	14-OCT-1999;	99US-0159330P.	
PR	14-OCT-1999;	99US-0159331P.	
PR	14-OCT-1999;	99US-0159637P.	
PR	18-OCT-1999;	99US-0159638P.	
PR	18-OCT-1999;	99US-0159584P.	
PR	21-OCT-1999;	99US-0160741P.	
PR	21-OCT-1999;	99US-0160767P.	
PR	21-OCT-1999;	99US-0160768P.	
PR	21-OCT-1999;	99US-0160770P.	
PR	21-OCT-1999;	99US-0160814P.	
PR	21-OCT-1999;	99US-0160815P.	
PR	22-OCT-1999;	99US-0160980P.	
PR	22-OCT-1999;	99US-0160981P.	
PR	22-OCT-1999;	99US-0160989P.	
PR	25-OCT-1999;	99US-0161404P.	
PR	25-OCT-1999;	99US-0161405P.	
PR	25-OCT-1999;	99US-0161406P.	
PR	26-OCT-1999;	99US-0161359P.	
PR	26-OCT-1999;	99US-0161360P.	
PR	26-OCT-1999;	99US-0161361P.	
PR	28-OCT-1999;	99US-0161920P.	
PR	28-OCT-1999;	99US-0161992P.	
PR	28-OCT-1999;	99US-0161993P.	
PR	29-OCT-1999;	99US-0162142P.	
Query Match 14.2%; Score 149; DB 3; Length 1403;			
Best Local Similarity 49.3%; Pred. No. 3e-36;			
Matches 389; Conservative 0; Mismatches 400; Indels 0; Gaps 0;			
Qy	177	TTTCTCCGCGTCTGCTGCTTGAAGACAAAGAACGACTCCATTTTTCCTCATGAA	236
Dd	412	TCTCTATCTCCCTTTGTTTCTCTGCTTCCACCTTTCTTCATCTCTTGTCTTT	471
Qy	237	ACCTCCTCTCTTTATCGCCGCTATCGTTGTTGTTGCTCGTGGGATTTGACAAATTACCT	296
Dd	472	CAAGACTCTGTTTGGATTATCTTTCTCTTGGTTTTCCTATTGCTATTGTTTAGACAATCTTTT	531

Qy	297	CTACTCTTACGGGTAGCTTATATCCCTGTTTCTACTGCGTCTTTGATCATCTCCGCGCA	356
Dd	532	ATACTCTTTTGGTCTTTTGTATCTCTGCTTCAACTTATTCGATTTCTTTGTTCTTCACA	591
Qy	357	ATTAGGCTTCACGTCTCTCTTTGTCATTTTATGGTGAAGCAAAAGTTACACCTTTTAC	416
Dd	592	GTTAGCTTTCAATGGTGTCTTCTTATTACATCAATCTCAGAAATCACTTGTTTGTAT	651
Qy	417	TATAACGCTATCGTTTGTCTCACTGGTGGCTAGTCCCTTGCCTTAACTCTGATAG	476
Dd	652	TCCTCTCTGTGTGTTGTTCTCTCTGCTCTCTGCTGTGTTGTTCTCTTACGATGATTC	711
Qy	477	TGACAGCTTGCACACAGACACACAGGAATATGTTGTTGGGTTCAATCATGACTCTGG	536
Dd	712	CAATAGCCCATCAGGAGATTCTAAGTGGAGTTACTTGAITGGTGTGTTTATGTACTGTTT	771
Qy	537	TGCAGCTCTTCTATGGGTTTATATGTCACCTTGTCCAGCTTCTTACAGAAATCTGG	596
Dd	772	TGCTTCTTATCTATCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	831
Qy	597	TCAGCGAATCACGTATACGCTCGCTCGCTCGAGTTCAGAGTGGTCTTATGCTTTGCTGCCAC	656
Dd	832	CAAGAGTGAGACGTTCTCTATGTTCTCGAGATGCAAACTCTATACGTCGCTTGTGGCTTC	891
Qy	657	TTGTGCTGCTCGTGGGATGCTAGCGCTGGGATTTCAAGGTGATAGCAGAGAAGC	716
Dd	892	TTGTGTTGCGGTTATCGGTTGTTGCGAGCGGGAATGATGTTGTGTGAGTGTGGAGAT	951
Qy	717	AAGAGATTTTAAAGCTTGGAGATCTTTGTACTATGTGGTGAATGTGTTCACGGCCATAAT	776
Dd	952	GGAGAGTTTTCACGAGGTCAGTCATTTATGTTTGTGTTGGTGGGACCGGTTTC	1011
Qy	777	CTGCAAGCAATTTTGTGGAGCTATTTGGTGTGATCTTCTGTGATGCTCTCTGTCTC	836
Dd	1012	GTGCAATTTGGTTCGTAGGAGCGGTGGCGCTATATTTCTGTTTCTTCTGCTGTTTTC	1071
Qy	837	TGGNATATGTCAGTCTCTGCTTCGGTGAGGATGATCTTGGCGCTCATTTGCTTTCCA	896
Dd	1072	GAACCTTATTTGGTACACTCTCACTGTTTACGCTCTTTCAGGCCAATTCGGGTGTTCCA	1131
Qy	897	GGAGAGTTTTCAGCGGGGAAAGGTGTCGCTTTTGGCTCTCTCCCTCTTGGGGATCAGTCTC	956
Dd	1132	CGACAGCTGACTGAGGTTAAGATGTCGCGATGCTCATTTGCCTTCATGCGGTTT	1191
Qy	957	TTATTTCTA 965	
Dd	1192	TTATATCTA 1200	
RESULT 11			
ADG88190			
ID	ADG88190 standard; cdna; 1047 BP.		
XX			
AC	ADG88190;		
XX			
Df	22-APR-2004 (first entry)		
XX			
DE	A. thaliana RPP4-upregulated pathogen infection-related gene #632.		
XX			
KW	Pathogen infection-related gene; plant; Peronospora parasitica;		
KW	defence mechanism; RPP4; pathogen resistance; transgenic plant; oomycete;		
KW	fungus; bacterium; virus; nematode; insect; aphid; gene; ss.		
XX			
OS	Arabidopsis thaliana.		
XX			
PN	WO200222675-A2.		
XX			
PD	21-MAR-2002.		
XX			
PF	14-SEP-2001; 2001WO-US028506.		
XX			
XX			
PR	15-SEP-2000; 2000US-0232778P.		





CC	thaliana PUP1 protein which is described in the method of the invention	
XX		
SQ	Sequence 1081 BP; 322 A; 231 C; 205 G; 323 T; 0 U; 0 Other;	
	Query Match 14.2%; Score 148.8; DB 3; Length 1081;	
	Best Local Similarity 50.6%; Pred. No. 3e-36;	
	Matches 360; Conservative 0; Mismatches 352; Indels 0; Gaps 0;	
QY	260 TCGTGTGTTGTTGTCGTGGGATTTGACATTAACCTCTACTCTTACGGGTAGCTTATA	319
DB	254 TAGTGTCTGGACTTCTTTAGGACAGATTCCTATCTGTACTCCATTGGACTTCTTACT	313
QY	320 TCCCTGTTCTACTGCGTCTTTTGATCATCTCGGCAATTAGGCTTCACTCTCTCTTGG	379
DB	314 TACCGTTTCTACCTATTCCCTGATCTGTGATCTCAGTTAGCCTTCAATGCTTCTTCT	373
QY	380 CATTTTATGGTGAAGCAAAAGTTTCACACCTTTCACTTATAAAGCTATCGTTTGTCTCA	439
DB	374 CTTATTTTCTTAACCTCACAAAACCTTACCCTATCATTTTAAATTCCTTTCTCCTTAA	433
QY	440 CTGGTGGTCGCTAGTCTTCCCTTAACCTGATAGTGAAGCTTGCACAAAGACAC	499
DB	434 CTATATCTTCCACCCCTACTTTCATTCATTAATGAGGACACAGACTCCACAAAAGTTACAA	493
QY	500 ACAAGGAATATGTTGTTGGTTCATCATGACTCTTGGTGACAGCTTCTCTATGGGTTTA	559
DB	494 AAGGAGATATGTCMAAGGTTTTCATGCACCGTTGCTGCTGCTGCTGTTATGGTCTAG	553
QY	560 TATTGCCACTTGTGAGCTTTCTTACAGAAATCTGFTCAGCAATCACGTATACGCTCG	619
DB	554 TCTTATCCCTACACAGCTAGCCTTCTTAAAGTCTTAAAGCAAAATTTCTCAGAAG	613
QY	620 CGCTCGAGTTCAGATGTTCTATGCTTGTGCGCACTTGTTGCTGCTGCTGCTGGGATGC	679
DB	614 TTATGGATATGATATCTACGTGAGTCTAGTGGCCAGTTGTGTAGCGTGGTGGGCTTT	673
QY	680 TAGCGCTGGCGATTTCAAGGTGATAGCAGGAGAAAGATTTTAAAGCTTCGAGAGT	739
DB	674 TTGCTAGCAGTGAAGTGGAACCTTTGACGATGAATGGATACTACAAACATGGGAAG	733
QY	740 CTTTGTACTATGTGGTGAATGTTTCACGGCCATAATCTGGCAAGCAATTTTGTGGGAG	799
DB	734 TATCTACATATTAACCTAGTGTGGACAGCTTTACCTGGCAGGATTTCTCCATCGGTG	793
QY	800 CTATTGGTTCATCTTCTGTCATCGTCTCTGGTCTCTGGAAATATGTCAGTCTCTGC	859
DB	794 GCACAGGACTGATCTTCGAGCTCTCCTCTCTATTCTCAAAATGCAATAGCGCTTTTGGGAC	853
QY	860 TTCGGTGACGGTGATCTTGGCCGTCATTTGCTTCCAGGAGAACTTTTCAGGCGGGAAAG	919
DB	854 TCCAGTGGTTCCTATCTTGGCTGTATCATTTTCCATGACAAATGAATGGGTTAAGG	913
QY	920 GTGTGGCTTGGCTCTCTCCCTCTGGGATCAGTCTCTTAATTTCTATGGACA	971
DB	914 TGAATTTCTATGATCTAGCTATTTGGGGTTTCACTTCTCTATGTTCTACCAACA	965
PN	EPI033405-A2.	
XX	FD	
XX	06-SEP-2000.	
XX	25-FEB-2000; 2000EP-00301439.	
PR	25-FEB-1999; 99US-0121825P.	
PR	05-MAR-1999; 99US-0123180P.	
PR	09-MAR-1999; 99US-0123548P.	
PR	23-MAR-1999; 99US-0125788P.	
PR	25-MAR-1999; 99US-0126264P.	
PR	29-MAR-1999; 99US-0126785P.	
PR	01-APR-1999; 99US-0127462P.	
PR	06-APR-1999; 99US-0128034P.	
PR	08-APR-1999; 99US-0128714P.	
PR	16-APR-1999; 99US-0129845P.	
PR	19-APR-1999; 99US-0130077P.	
PR	21-APR-1999; 99US-0130449P.	
PR	23-APR-1999; 99US-0130510P.	
PR	23-APR-1999; 99US-0130891P.	
PR	28-APR-1999; 99US-0131449P.	
PR	30-APR-1999; 99US-0132048P.	
PR	30-APR-1999; 99US-0132407P.	
PR	04-MAY-1999; 99US-0132484P.	
PR	05-MAY-1999; 99US-0132485P.	
PR	06-MAY-1999; 99US-0132486P.	
PR	07-MAY-1999; 99US-0132487P.	
PR	06-MAY-1999; 99US-0132863P.	
PR	11-MAY-1999; 99US-0134256P.	
PR	14-MAY-1999; 99US-0134218P.	
PR	14-MAY-1999; 99US-0134219P.	
PR	14-MAY-1999; 99US-0134221P.	
PR	14-MAY-1999; 99US-0134370P.	
PR	18-MAY-1999; 99US-0134768P.	
PR	19-MAY-1999; 99US-0134941P.	
PR	20-MAY-1999; 99US-0135124P.	
PR	21-MAY-1999; 99US-0135353P.	
PR	24-MAY-1999; 99US-0135629P.	
PR	25-MAY-1999; 99US-0136021P.	
PR	27-MAY-1999; 99US-0136392P.	
PR	28-MAY-1999; 99US-0136782P.	
PR	01-JUN-1999; 99US-0137222P.	
PR	03-JUN-1999; 99US-0137528P.	
PR	04-JUN-1999; 99US-0137502P.	
PR	07-JUN-1999; 99US-0137724P.	
PR	08-JUN-1999; 99US-0138094P.	
PR	10-JUN-1999; 99US-0138540P.	
PR	14-JUN-1999; 99US-0138847P.	
PR	16-JUN-1999; 99US-0139452P.	
PR	16-JUN-1999; 99US-0139453P.	
PR	17-JUN-1999; 99US-0139492P.	
PR	18-JUN-1999; 99US-0139454P.	
PR	18-JUN-1999; 99US-0139455P.	
PR	18-JUN-1999; 99US-0139456P.	
PR	18-JUN-1999; 99US-0139457P.	
PR	18-JUN-1999; 99US-0139458P.	
PR	18-JUN-1999; 99US-0139459P.	
PR	18-JUN-1999; 99US-0139460P.	
PR	18-JUN-1999; 99US-0139461P.	
PR	18-JUN-1999; 99US-0139462P.	
PR	18-JUN-1999; 99US-0139463P.	
PR	18-JUN-1999; 99US-0139750P.	
PR	18-JUN-1999; 99US-0139763P.	
PR	21-JUN-1999; 99US-0139817P.	
PR	22-JUN-1999; 99US-0139899P.	
PR	23-JUN-1999; 99US-0140353P.	
PR	23-JUN-1999; 99US-0140354P.	
PR	24-JUN-1999; 99US-0140895P.	
PR	28-JUN-1999; 99US-0140823P.	
PR	29-JUN-1999; 99US-0140991P.	
PR	30-JUN-1999; 99US-0141287P.	
PR	01-JUL-1999; 99US-0141842P.	

RESULT 13  
AAC47974  
ID AAC47974 standard; DNA; 2660 BP.  
XX  
XX AAC47974;  
AC  
XX  
DT 18-OCT-2000 (first entry)  
XX  
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 55799.  
DE  
XX  
XX Hybridisation assay; genetic mapping; gene expression control;  
KW protein identification; signal transduction pathway; metabolic pathway;  
KW promoter; termination sequence; ss.  
XX  
OS Arabidopsis thaliana.  
XX

PR	01-JUL-1999;	99US-0142154P.
PR	02-JUL-1999;	99US-0142055P.
PR	06-JUL-1999;	99US-0142390P.
PR	08-JUL-1999;	99US-0142803P.
PR	09-JUL-1999;	99US-0142920P.
PR	12-JUL-1999;	99US-0142977P.
PR	13-JUL-1999;	99US-0143342P.
PR	14-JUL-1999;	99US-0143324P.
PR	15-JUL-1999;	99US-0144005P.
PR	16-JUL-1999;	99US-0144085P.
PR	19-JUL-1999;	99US-0144325P.
PR	19-JUL-1999;	99US-0144331P.
PR	19-JUL-1999;	99US-0144332P.
PR	19-JUL-1999;	99US-0144333P.
PR	19-JUL-1999;	99US-0144334P.
PR	20-JUL-1999;	99US-0144352P.
PR	20-JUL-1999;	99US-0144632P.
PR	20-JUL-1999;	99US-0144884P.
PR	21-JUL-1999;	99US-0144814P.
PR	21-JUL-1999;	99US-0145086P.
PR	21-JUL-1999;	99US-0145085P.
PR	22-JUL-1999;	99US-0145087P.
PR	22-JUL-1999;	99US-0145192P.
PR	22-JUL-1999;	99US-0145145P.
PR	23-JUL-1999;	99US-0145218P.
PR	23-JUL-1999;	99US-0145224P.
PR	26-JUL-1999;	99US-0145276P.
PR	27-JUL-1999;	99US-0145913P.
PR	27-JUL-1999;	99US-0145918P.
PR	27-JUL-1999;	99US-0145919P.
PR	28-JUL-1999;	99US-0145951P.
PR	02-AUG-1999;	99US-0146386P.
PR	02-AUG-1999;	99US-0146388P.
PR	03-AUG-1999;	99US-0146389P.
PR	04-AUG-1999;	99US-0147204P.
PR	04-AUG-1999;	99US-0147302P.
PR	05-AUG-1999;	99US-0147192P.
PR	06-AUG-1999;	99US-0147260P.
PR	06-AUG-1999;	99US-0147303P.
PR	09-AUG-1999;	99US-0147416P.
PR	09-AUG-1999;	99US-0147493P.
PR	09-AUG-1999;	99US-0147935P.
PR	10-AUG-1999;	99US-0148171P.
PR	11-AUG-1999;	99US-0148319P.
PR	12-AUG-1999;	99US-0148341P.
PR	13-AUG-1999;	99US-0148565P.
PR	13-AUG-1999;	99US-0148684P.
PR	16-AUG-1999;	99US-0149368P.
PR	17-AUG-1999;	99US-0149175P.
PR	18-AUG-1999;	99US-0149426P.
PR	20-AUG-1999;	99US-0149722P.
PR	20-AUG-1999;	99US-0149723P.
PR	20-AUG-1999;	99US-0149929P.
PR	23-AUG-1999;	99US-0149902P.
PR	23-AUG-1999;	99US-0149930P.
PR	25-AUG-1999;	99US-0150566P.
PR	26-AUG-1999;	99US-0150884P.
PR	27-AUG-1999;	99US-0151065P.
PR	27-AUG-1999;	99US-0151066P.
PR	27-AUG-1999;	99US-0151080P.
PR	30-AUG-1999;	99US-0151303P.
PR	31-AUG-1999;	99US-0151438P.
PR	01-SEP-1999;	99US-0151930P.
PR	07-SEP-1999;	99US-0152363P.
PR	10-SEP-1999;	99US-0153070P.
PR	13-SEP-1999;	99US-0153758P.
PR	15-SEP-1999;	99US-0154018P.
PR	16-SEP-1999;	99US-0154039P.
PR	20-SEP-1999;	99US-0154779P.
PR	22-SEP-1999;	99US-0155139P.
PR	23-SEP-1999;	99US-0155486P.
PR	24-SEP-1999;	99US-0155659P.
PR	28-SEP-1999;	99US-0156458P.
PR	29-SEP-1999;	99US-0156596P.
PR	04-OCT-1999;	99US-0157117P.
PR	05-OCT-1999;	99US-0157533P.
PR	06-OCT-1999;	99US-0157865P.
PR	07-OCT-1999;	99US-0158029P.
PR	08-OCT-1999;	99US-0158232P.
PR	12-OCT-1999;	99US-0158369P.
PR	13-OCT-1999;	99US-0159293P.
PR	13-OCT-1999;	99US-0159294P.
PR	13-OCT-1999;	99US-0159295P.
PR	14-OCT-1999;	99US-0159329P.
PR	14-OCT-1999;	99US-0159330P.
PR	14-OCT-1999;	99US-0159331P.
PR	14-OCT-1999;	99US-0159637P.
PR	18-OCT-1999;	99US-0159584P.
PR	21-OCT-1999;	99US-0160741P.
PR	21-OCT-1999;	99US-0160767P.
PR	21-OCT-1999;	99US-0160768P.
PR	21-OCT-1999;	99US-0160770P.
PR	21-OCT-1999;	99US-0160814P.
PR	21-OCT-1999;	99US-0160815P.
PR	22-OCT-1999;	99US-0160980P.
PR	22-OCT-1999;	99US-0160981P.
PR	22-OCT-1999;	99US-0160989P.
PR	25-OCT-1999;	99US-0161404P.
PR	25-OCT-1999;	99US-0161405P.
PR	25-OCT-1999;	99US-0161406P.
PR	26-OCT-1999;	99US-0161359P.
PR	26-OCT-1999;	99US-0161360P.
PR	26-OCT-1999;	99US-0161361P.
PR	28-OCT-1999;	99US-0161920P.
PR	28-OCT-1999;	99US-0161992P.
PR	28-OCT-1999;	99US-0161993P.
PR	29-OCT-1999;	99US-0162142P.
<div>Query Match 13.6%; Score 142.2; DB 3; Length 2660; Best Local Similarity 48.7%; Pred. No. 6.4e-34; Matches 462; Conservative 0; Mismatches 468; Indels 19; Gaps 2;</div>		
QY	19	GTTCTTGTAATCATAACTGTATATCTTGCCATTGGAACCTGGAGGCCCTCTAATG 78
DB	197	GTTCCTGATCCGTTAACATCTTTTCTTATTTGGTCAAGCTGCTTCGGTTCTTCTT 256
QY	79	ATGCGTCTCTACTTCCAAATGGTGGCGAAAGGATCTGTTTCCAAGCTTCCTTCAAACC 138
DB	257	GGTAGGTTTACTATGATGAAGGTGGAACACAGTAATGATGGCACTCTTGTTCNAAG 316
QY	139	GTGCTGTCACATCTTCTCCCTCTTCTCTTA--TCTTTCCTCCGCGTCTGCTGTT 196
DB	317	GCTGCTTTCTCTACTCTATATCCGCTTTTGTCTACTTCCGCTCTTCGCTAGTAGAG 376
QY	197	GCCTTGAGAACAGAACGACTCCATTTTCTCATGAACCTCTCTCTTATCGCG 256
DB	377	TCTTGGAGAGTTCGTGTTCACTCAAGTACA-----TTGTTTGTATCT 419
QY	257	CTATCGTTGTTGGTTTGTCTCGTGGATTTTGACAATTACTCTACTCTTTACGGTTAGCTT 316
DB	420	ATGTTTGTCTCGGTGTGATCACTGCTGGAGATAATATGTTATCTCTCTGTGACTTTTGT 479
QY	317	ATATCCCTGTTTCTACTGCGTCTTTTGATCATCTCGCGCAATTAGGTTCTACTGCTCTCT 376
DB	480	ACCTCTGTCATCGAGGTATTTCGCTCACTTTGCGCTACTCAGTACTGTTTCAACGGGTGT 539
QY	377	TTCATTTTATGTTGAGCAAAAGTTCAACCTTTCACTATATAACCTATCGTTTTCG 436
DB	540	TCTCTTATTTCATCAATGCTCAGAAGTTCACTGCTTTGATTTCTCACTCCGTTGTTCTCT 599



Db 890 AAGCTACACCTTTGGTCTTGTGGGATGTAATTCGAGTCAATCCCTCTGTCTCCAAATT 949  
Qy 842 TTATGGTCAGTCTCTGCTTCCGGTGACGGTGATCTTGGCGGTCATTTGCTCCAGGAGA 901  
Db 950 CCATAACAGCTGTGGGATTCGCTATAGTCCAGTTCGGGCGAGTAGATTTTCCAGATA 1009  
Qy 902 AGTTTCAGGCGGGGAAAGGTGTCGCTTTGGCTCTCTCCCTCTGGGGATCAGTCTCTTATT 961  
Db 1010 GAATGGACGATCAAAAATCTTCTCCATTAATTTTAGCTATCTGCGGCTTCCCTTCATTCG 1069  
Qy 962 TCTATGACAGGTTAAATCCGAGGAGAACTAAAGGCTCAGGATACA 1008  
Db 1070 TCTATCAGCACTACCTCGACGAAAGATGTAATACTAGCCACACA 1116

RESULT 15  
ABN85767  
ID ABN85767 standard; cDNA; 83698 BP.  
XX AC ABN85767;  
XX 21-OCT-2002 (first entry)  
XX Arabidopsis yellow stripe1-like 4 encoding cDNA SEQ ID NO 9.  
XX Maize; transgenic; plant; yellow stripe1-like; ysl; Arabidopsis;  
KW iron uptake; bioremediation; yellow stripe 1; ysl; gene; ss.  
XX Arabidopsis sp.  
XX WO200240688-A2.  
XX 23-MAY-2002.  
XX 16-NOV-2001; 2001WO-US043101.  
XX 16-NOV-2000; 2000US-0249222P.  
XX (UYVA ) UNIV YALE.  
XX Walker EL, Dellaporta S;  
XX WPI; 2002-490144/52.  
XX P-PSDB; AB883920.  
XX New yellow stripe1 and yellow stripe1-like genes, useful for altering the  
XX distribution of iron within the plant body so that edible parts of crop  
XX plants have more iron, or for producing plants useful in enhancing iron  
XX uptake from soil.  
XX Claim 1; Page 114-159; 187pp; English.

XX The invention relates to an isolated nucleic acid molecule (I), maize  
XX yellow stripe 1 (ysl) or yellow stripe1-like (ysl) from Arabidopsis  
XX (ABN85763-ABN85771). (I) is useful for generating transgenic plants which  
XX can be used for enhancing iron uptake from soil and for bioremediation of  
XX metal or heavy metal contaminated soil. (I) may also be used to alter the  
XX distribution of iron within the plant body so that edible parts of crop  
XX plants have more iron. Transgenic plants may also be used in conventional  
XX plant breeding schemes to produce progeny which also contain the gene of  
XX interest. The present sequence is that of the Arabidopsis ysl encoding  
XX cDNA of the invention  
XX  
XX Sequence 83698 BP; 26295 A; 15150 C; 15104 G; 27149 T; 0 U; 0 Other;

Query Match 12.7%; Score 133; DB 6; Length 83698;  
Best Local Similarity 48.0%; Pred. No. 3.8e-30;  
Matches 455; Conservative 0; Mismatches 475; Indels 17; Gaps 2;  
Qy 20 TTCTGTATATCACTAATCTATATTTCTGGCCATTGGAACTGTGGAGGCCCTCTAATGA 79  
Db 67235 TTCTGTGTTTCAATAGCATCTTTCTCTCATCTCTGCTCAAGCCATTTCTGTCTCTTG 67294

Qy 80 TGCGTCTCTACTTCCAAAATGGTGGCAAGGATCTGTTTCCAAAGCTTCTCTTCAAACCG 139  
Db 67295 GTCCGCTTTTATTACAAATGAAGGTGAAACAGTAATGGATCTCTACTCTTGTTCAAACTG 67354  
Qy 140 TTGG-TTGTCCACTCAATTTTCTTCCCTCTCTCTTATCTTTCTCGGCGGTGCTGTGC 198  
Db 67355 GTGGCTTTCCCAATCTTTATCTCCCTCTTCTCTCTCTCTCTTCCAAATCTTCTCTT 67414  
Qy 199 CTTGAAGAAACAAGAAACGACTCCATTTTCTCATGAAACCTCTCTCTTATTCGCGGT 258  
Db 67415 CTT-----CTTCTTCTTCTTCTTCTTCAAGACTCTGGTTTGGATTAT 67458  
Qy 259 ATCGTTGTGGTTGCTCGTGGGATTTGACAATTAACCTCTACTCTTACGGGTTAGCTTAT 318  
Db 67459 CTTTCTCTGGTTTGTCTATTGGTTAGATAAATTTTATATACTCTGTGGACTTTTGTAT 67518  
Qy 319 ATCCCTGTTTCTACTGCGCTTTTGATCATCTCCGGCGCAATTAGGCTTCACTGCTCTCTT 378  
Db 67519 CTTCTCTGCTTCAACTTATTCGAATCTATGTGCTTCCACAGTTAGCTTTCAATGGTCTCT 67578  
Qy 379 GCATTTTATGGTGAAGCAAAAGTTACACCTTTTCACTATATAAAGCTATCGTTTGTCT 438  
Db 67579 TATTAATACATCAATCTCAGAAAATCACTGTGTTGATTTCTTCTCAGTGTGTTCTC 67638  
Qy 439 ACTGGTGGTGGCTAGTCTTGGCCCTTAACTCTGTAGTGACAAGCTTGCAAAACGAGACA 498  
Db 67639 TCTATCTCTGCTGTGTTGCTTCTTCTGACGATGATTCAAATAGCCCATCAGGAGATTCT 67698  
Qy 499 CACAAGGAATATGTTGGTGGTTTCATCATGACTCTTGGTGACGCTCTCTCTATGGGTTT 558  
Db 67699 AAGTGGAGTTACTTGAATGGGTGTTTCTGTGCAGTTTTTGTCTCTTATCTATCTCT 67758  
Qy 559 ATATTGCCACTTGTGCGAGCTTTCTTACAAGAAATCTGGTTCAGCGAATCACGPTATACGCTC 618  
Db 67759 CAGCTCTCTCTATGACAGTTTTTCTTTCGAGAAGGTTCTCAAGAGTGAGACTCTCTCTATG 67818  
Qy 619 GCGCTCGAGTTCAGATGCTCTTATGCTTTGCTGCCACTTGTGTCTGCTCGTGGGGATG 678  
Db 67819 GTTCTCGAGATGCAAAATCTATAGCTGCTTGTGGCTTCTTGTGTAGCGGTTATCGGATTG 67878  
Qy 679 CTAGCCGTGGGGATTTCAAGGTGATAGCAGGAGAACAGAGATTTTAAGCTTGGAGAG 738  
Db 67879 TTCCGAAGCGGGGAATGGATGTTGTGAGTGTGGAGATGGAAGAGTTTTCAGGAAGTCAA 67938  
Qy 739 TCTTTGTACTATGTTGTTGTTGTTCACGGCCATAATCTGCAAGCAATTTTTCGTGGGA 798  
Db 67939 GTCAATTAATGTTTGAATTTGCTGGGGCAGCGGTTTCGTGTCATTTGGGTTGTAGGA 67998  
Qy 799 GCTATTTGGGTTGATCTTCTGTGCATCGTCTCTGGTCTCTGGAAATTAATGTCAGTCTCTG 858  
Db 67999 GCGGTGTCGCTTATATTTCTGTTCTGCTGTTTTCAAAACCTTATTTAGTACGCTCTCA 68058  
Qy 859 CTTCCGGTGACGGTGATCTTGGCGCTCATTTGCTTCCAGGAGAGATTTTCAGGCGGGGAAA 918  
Db 68059 CTCATTTGTACGCTCTCGCGGCCAATTGCGGTGTTTCCATGACAGCTGACTGAGGTTAAG 68118  
Qy 919 GGTGTCGCTTTGGCTCTCTCCCTCTGCGGATCAGTCTCTTATTTCTTA 965  
Db 68119 ATGTTGGCGATGCCCATCGCCTTCACAGGATTCAGGTTTATATATCTA 68165

Search completed: November 1, 2004, 07:19:42  
Job time : 559.844 secs

**THIS PAGE BLANK (USPTO)**

Result No.	Score	Query			DB	ID	Description
		Match	Length				
c	1	1049	100.0	1049	6	AX033545	Sequence
	2	997.4	95.1	1075	8	AF078532	Arabidops
	3	701.8	66.9	114144	8	U78721	Arabidops
	4	528	50.3	1225	8	AX033544	Sequence
	5	528	50.3	1227	8	AF078531	Arabidops
	6	526.4	50.2	1250	8	AF17664	Arabidops
	7	519	49.5	1102	8	BT005504	Arabidops
	8	478	45.6	1067	6	AX033553	Sequence
	9	408.2	38.9	121720	8	AC021044	Arabidops
	10	251.6	24.0	1505	8	AK066610	Oryza sat
	11	242	23.1	1458	8	AK061667	Oryza sat
	12	226	21.5	1363	8	AK064135	Oryza sat
	13	167.6	16.0	1165	8	AY036558	Arabidops
	14	167.6	16.0	1428	8	AY074546	Arabidops
	15	167.6	16.0	3387	6	AX057517	Sequence
	16	167.6	16.0	3387	6	AX652015	Sequence
	17	167.6	16.0	82697	8	ATT9A21	Sequence
	18	167.6	16.0	194143	8	AL021713	Arabidops
	19	163.6	15.6	73179	8	AL161548	Arabidops
	20	163.6	15.6	73179	8	AP006383	Lotus cor







mRNA  
CDS  
repeat\_region  
gene  
mRNA  
CDS  
repeat\_region  
repeat\_region  
gene  
mRNA  
CDS  
gene  
mRNA  
CDS

join(<7252..7434,7522..7713,7864..8069,8275..9045)  
/gene="At2g33700"  
join(7252..7434,7522..7713,7864..8069,8275..9045)  
/gene="At2g33700"  
/codon\_start=1  
/product="putative protein phosphatase 2C"  
/protein\_id="AAC69126.1"  
/db\_xref="GI:1707015"  
/translation="MSMDFSLLTLLEGDNFKONTSSATEIDTLENLDDTROISKGKP  
PHLTSATRIQLAANADVDVNLVMSKSLDDKSEFLPVYRSGSCAEQCAQFWEDEHI  
CIDLVNHLGAAGAACSSILGAFYGVDFDGHGGTDAAHFVKNIILRIVEDSFPPLCVKKA  
IKSAFLADYEFADDSSLDLSGGTALTARIFGRLLIIANAGDCRAVLGRGRIELS  
KDHPCNTAEKRIEKGGLGVYDGLNGQLSVARAIGDWMHKKPGKSACPLPEPELQ  
ETDUSEDDFIMCGDGLVDMVMSQCAVTIARKELMHNDPERCSRSLREALKRNTC  
DNLTIVVCFSPDPFQRIETRMQSRVRSISAEGLNLLKGLVDGP"  
complement(11621..11649)  
/rpt\_family="AT rich"  
complement(12323..13704)  
/gene="At2g33710"  
note="synonym: TlB8.3"  
complement(join(<12323..12798,13524..>13704))  
/gene="At2g33710"  
complement(join(12323..12798,13524..13704))  
/gene="At2g33710"  
/codon\_start=1  
/product="putative AP2 domain transcription factor"  
/protein\_id="AAC69127.1"  
/db\_xref="GI:1707016"  
/translation="MHSGKRPLSPESMAGNREKKELCCSTLSESDVDFVSELTCQ  
PIPSIDDOSSSLTLOKSNRSORNYRGVQRPMGWAAIRIDPNKAARWLGTFTDQA  
EEAALVDKAAFERGHAKLNFFEHIRVAPTQLYPSPATSHDRILVTPSPPPPIAP  
DILDDQYGHFOSRSSDSANLNMNLSSSSSSLNHQGLRNLNEDGENVKNISIHKKRK  
"  
complement(14124..14156)  
/rpt\_family="AT rich"  
complement(14389..14425)  
/rpt\_family="(TA)n"  
16828..18041  
/gene="At2g33720"  
note="synonym: TlB8.29; predicted by gensecan and  
genefinder"  
join(<16828..17139,17373..>18041)  
/gene="At2g33720"  
join(16828..17139,17373..18041)  
/gene="At2g33720"  
/codon\_start=1  
/product="hypothetical protein"  
/protein\_id="AAC69145.1"  
/db\_xref="GI:2253580"  
/translation="MKRPPPSASKTQYLFDHDESSNSKSLVSTTLSSLSCENPNKR  
KNSDEVNLISCIPRDYKLTQVERKIAIRNRLNSYOEKADEWYGVSTELTLFKDPWII  
KKVHFASVLDMAFNSVSNTHCLDTPESPENAEESLVSLDLCFYDKTWPDPNVAYNK  
PTSEAINLAWMTMSKRAREEKXYVSTELTLLTVADPWLKAMTKSSIGNLYRL  
VLKASFDVIDIHLRLPLDDOMMKVEDSLAVEVYDHTDTSVHNIALAKWAKSSSFVLV  
SGWPKCFVDREGLQVDVIGMYDRSBSKLHFCVLSRSETMDSAPLPSP"  
complement(18731..20932)  
/gene="At2g33730"  
note="synonym: TlB8.4"  
complement(<18731..>20932)  
/gene="At2g33730"  
complement(18731..20932)  
/gene="At2g33730"  
/codon\_start=1  
/product="putative U5 small nuclear ribonucleoprotein, an  
RNA helicase"  
/protein\_id="AAC69128.1"  
/db\_xref="GI:1707017"  
/translation="MKRTIINEVSGTSVNTLDDLSIAKKPVFLTKKQREELAKRQDQ  
ISEQRVREQLGRPEPETEDVNSGDTNRDKDRDRDRDRDRDRDRDRDRDRDRDR  
DR  
DAIKEQYLGKKPKKRVIRPSEKFRFDFMENTEDTSRDNMVLYQNPEHAQLLFGRF  
RAGMDREOKQAAKHEKMERDEIRKKDGI VEKPEEAAAORVEEAAADTVDSFDMRVYD  
RHWSDKRLBETWDRIFREDFNI SYKSGRI PRPMRSWEESKLITSELLKAVEGAGYK

KPSPIOMAAIPLGLQORDVIGIAETSGKTAFAVLPMLAVLSRLPMPMSENETEGPYA  
VWNAFRELAAQIEEBETKFAHYLGRVTSIVGGQSTIEBOGLKITQCEBVIATPGRLL  
IDCLERYAVLNQCNVYVLDDEADMDMGFEPPQVGLDAMPSSNLKPENEEBELDK  
KIYRTTYMFSAITPMPGVERLARKYLNRNPPVVTITGAKTTLDSIQHVIIMKSEKFR  
KYNLDELDEKTAIVFVNTKKNDSIAKNLDKAGYRVTLTHGGKSQBQRISLSEGPRA  
KYNLVLAVTIVVGRGIDIPDAHVINYDMPKHIEMVTHRIGRTGRAGKSGVATSLFLTL  
HDTVEFYDLKQMLVQSNASVPPELARHEASRFKPGTVPDRPPHSDTVIYN"  
complement(20588..20718)  
/rpt\_family="(CAGAGA)n"  
complement(21321..22493)  
/gene="At2g33735"  
/notes="synonym: TlB8.17; supported by full length cDNA:  
Ceres: 269036"  
complement(join(<21321..21655,21744..21814,22051..22126,  
22221..22298,22384..>22493))  
/gene="At2g33735"  
complement(join(21584..21655,21744..21814,22051..22126,  
22221..22298,22384..22446))  
/gene="At2g33735"  
/codon\_start=1  
/product="Expressed protein"  
/protein\_id="AAM15518.1"  
/db\_xref="GI:20198310"  
/translation="MMSDWDGSHGDFDFEDYKDHYKVLNCDASDDEIRSSPFR  
LALKWHPDFKPEEDSATSRPQENAYQVLSDFIARQEDYDKRMRIYEHNTLLNEY  
LKLILITCNGLGVNHYLM"  
22524..22544  
repeat\_region  
Query Match 66.9%; Score 701.8; DB 8; Length 114144;  
Best Local Similarity 99.7%; Pred. No. 1.9e-199;  
Matches 703; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 AAGATGAAGATGAAGACAGTTCCTGTAATCATAACTGTATATCTTGGCCATTGGAAC 60  
Db 26320 AAGATGAAGATGAAGACAGTTCCTGTAATCATAACTGTATATCTTGGCCATTGGAAC 26261  
Qy 61 TGTGGAGGCCCTCTAATGATGCGTCTCTACTCCAAATCGTGGCGAAGATCTGTTT 120  
Db 26260 TGTGGAGGCCCTCTAATGATGCGTCTCTACTCCAAATCGTGGCGAAGATCTGTTT 26201  
Qy 121 CCAAGCTTCCTTCAAAACCGTTGGTTGTCCACTCATTTCTTCCCTCTCTTATCTTTC 180  
Db 26200 CCAAGCTTCCTTCAAAACCGTTGGTTGTCCACTCATTTCTTCCCTCTCTTATCTTC 26141  
Qy 181 CTCGCGCGTCTGCTGCTTGAAGAACAAAGACGATCCATTTTCTCATGAACACCT 240  
Db 26140 CTCGCGCGTCTGCTGCTTGAAGAACAAAGACGATCCATTTTCTCATGAACACCT 26081  
Qy 241 CCTCTCTTTATCCCGCTATCGTTGTTGGTTTGTCTGTGGATTTGACAATTAACCTCTAC 300  
Db 26080 CCTCTCTTTATCCCGCTATCGTTGTTGGTTTGTCTGTGGATTTGACAATTAACCTCTAC 26021  
Qy 301 TCTTTACGGGTAGCTTATATCCCTGTTTCTTACTCGCTCTTTTGATCATCTCGCGCAATTA 360  
Db 26020 TCTTTACGGGTAGCTTATATCCCTGTTTCTTACTCGCTCTTTTGATCATCTCGCGCAATTA 25961  
Qy 361 GGCTTCACTGCTCTCTTTTGATATTTTATGGTGAAGAAAGTTTACACCTTTTACATATA 420  
Db 25960 GGCTTCACTGCTCTCTTTTGATATTTTATGGTGAAGAAAGTTTACACCTTTTACATATA 25901  
Qy 421 AACGCTATCGTTTGTGCTCACTGGTGGCGGTAGTCTTCCCTTAACCTCTGATAGTAC 480  
Db 25900 AACGCTATCGTTTGTGCTCACTGGTGGCGGTAGTCTTCCCTTAACCTCTGATAGTAC 25841  
Qy 481 AAGCTTCAAAACGAGACACAAAGGAATATGTTGTTGGGTTTCATCATGACTCTTTGGTGA 540  
Db 25840 AAGCTTCAAAACGAGACACAAAGGAATATGTTGTTGGGTTTCATCATGACTCTTTGGTGA 25781  
Qy 541 GCTCTTCTATGGGTTTATATGCGCACTTGTGAGCTTTCTTACAAGAAATCTGCTGAC 600  
Db 25780 GCTCTTCTATGGGTTTATATGCGCACTTGTGAGCTTTCTTACAAGAAATCTGCTGAC 25721  
Qy 601 CGAATCACGTATACGCTCGCGCTCGAGTTCACAGATGCTCTTATGCTTTTGTGCGCACTTGT

Db	25720	CGAATCAGTATACGCTCGCGCTCGAGTTCAGATGCTTATATGCTTGTGCGCACTTGT	25661
Qy	661	GTCTGCTCGTGGGATGCTAGCGCTGGCGATTTCAAGGTGATA	705
Db	25660	GTCTGCTCGTGGGATGCTAGCGCTGGCGATTTCAAGGTAAAA	25616
RESULT 4			
AX033544			
LOCUS	AX033544	1225 bp	DNA linear PAT 21-SEP-2000
DEFINITION	Sequence 1 from Patent WO049152.		
ACCESSION	AX033544		
VERSION	AX033544.1	GI:10280288	
KEYWORDS			
SOURCE	Arabidopsis thaliana (thale cress)		
ORGANISM	Arabidopsis thaliana		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.		
AUTHORS	Andre, B., Buerkle, L., Frommer, W.B. and Gillissen, B.		
TITLE	Nucleic acids that code for a nucleobase transporter		
JOURNAL	Patent: WO 0049152-A 1 24-AUG-2000;		
	ANDRE BRUNO (BE) ; BUERKLE LUKAS (DE) ; FROMMER WOLF B (DE) ;		
	GILLISSEN BERND (DE)		
FEATURES	Location/Qualifiers		
source	1..1225		
ORIGIN	/organism="Arabidopsis thaliana"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:3702"		
Query Match	50.3%;	Score 528;	DB 6; Length 1225;
Best Local Similarity	70.8%;	Pred. No. 2.2e-147;	
Matches	762;	Conservative 0;	Mismatches 285; Indels 33; Gaps 3;
Qy	1	AAGATGAAGATGAAGACAGTCTTGTAAATCATAACTGTATATCTTTGGCCATGGAAAC	60
Db	16	AAGAAGAGATGAAGATGGTTTGATATATCATAACTGTATATCTCTACTATAGGAACA	75
Qy	61	TGTGGAGCCCTCTAATGATGCTCTACTTCCAAATGGTGGGAAAGGATCTGGTTT	120
Db	76	TGTGGAGGTCCTTTGTAACTCTCTACTTCCAAATGGCGGAAACGAATCTGGTTC	135
Qy	121	CCAAGTCTCTCAACCGTGGTGTGCCACTATTTCTTCCCTCTCTCTATCTTTC	180
Db	136	ATGAGCTTCTATCAACCGTGGTGTTCCAATCATCTCATCCCTCTCTTGGTCTCTTC	195
Qy	181	CTCCGCGCTCGTGGTTG-----CCTTGAAGAACAAAGAAACGACTCCA	222
Db	196	CTCAGCGTCCGCGGCAACCCCTAACACGCGGAAACAAAGCGGAAACAAAG	255
Qy	223	TTTTTCTCATGAACACTCTCTCTTTATCGCGCTATCGTTGGTTGCTCGTGGGA	282
Db	256	CTCTTCTCATGAAACCTCTCTGTTTATCGCTCCATTTGTCATAGGGTTGCTCACAGGA	315
Qy	283	TTTGACAATTTACTCTACTCTTACGGGTAGCTATATCCCTGTTTCTACTCGTCTTTC	342
Db	316	CTTGACAATCTATATATTTTACGGATATAGCATATCTGCCAGTTTCAACTTCATCGCTC	375
Qy	343	ATCATCTCCGCGCAATTAGCGTTTCACTGCTCTCTTTGCAATTTTATGGTGAAGCAAAAG	402
Db	376	ATAATCGGAACCTCACTAGCTTTCAACGCTCTCTTTCGCTTTCTTGTAGTCAAGCAAAAG	435
Qy	403	TTCAACCTTTCTACTATAAAGCTATCGTTTTCCTCACTGGTGGTCCGCTGCTTGC	462
Db	436	TTCACTCCGTTCTCCATAAAGCGGTCGTTTGTGACGGTTGGTATCGGATCTCTGCG	495
Qy	463	CTTAACTCTGATAGTACAGCTTGCAAACGACACACACAGGAATATGTTGGGTTC	522
Db	496	TTACACAGTGTGAGACAAACCGGCTAAGGAGCAAGAAAGAGTATGTTGGTGGTTC	555
Qy	523	ATCATGACTTCTGGTGCAGCTCTTCTCTATGGGTTTATATGCCACTTGTGCGAGCTTCT	582
Db	556	TTGATGACTGTGTTGCAGCTCTTCTCTATGCTTTTATATATACCGCTCGTTGAGCTAACT	615
Qy	583	TACAAGAATCTGGTCAGCGAATCAGTATACGCTCGCGCTCGAGTTCAGATGGTCTTAA	642
Db	616	TACAAGAAAGCTCGTCAAGAAATCACTTTCCACATTTGCTTGAGATTGAGATGGTCAATG	675
Qy	643	TGCTTTGCTGCCACTTGTGCTGCGCTCGTGGGATGCTAGCGCTCGCGATTTTCAAGGTG	702
Db	676	TGCTTTGCTGCTACTTTTCTGTCATTTGTCATTTGTCATGTTGAGATTTTAAGGTG	735
Qy	703	ATAGCAGGAGAAAGCAAGAGATTTTAAAGCTTGAGAGTC---TTTGTACTATGTGGTGATT	759
Db	736	ATAGCAAGAGAAAGCAAGAGATTCAAGATTGAGGATCAGTGTGTTTACTATGCAATTGATA	795
Qy	760	GTGTTTACGCGCCATAATCTGGCAAGCATTTTGTGGGAGCTATTGGGTTGATCTTCTCT	819
Db	796	GTGATCACAGGAATAATATGGCAAGGTTTCTTCTTAGGAGCCATAGGGATTTGTTTGT	855
Qy	820	GCATCGTCTCTGCTCTCTGGAATTATGTCAGTGTCTCTCTCCGCTGACGGTGATCTTGT	879
Db	856	GCATCATCATCTAGTCTCTGCTGTTCTGATAAGTGTCTGCTTCCGGTGACTGAAGTTTTC	915
Qy	880	GCGTCATTTGCTTCCAGGAGAAAGTTTTCAGCGGGGAAAGGTGTCGCTTTGGCTCTCTCC	939
Db	916	GCGTCTGTTTGTTCGGGAGAAAGTTTTCAGGCAGAGAAAGGTGCTCTCTCTCTCTTCT	975
Qy	940	CTCTGGGATCAGTCTCTTATTTCTATGACAGAGTTAAATCC-----GAGGAG	987
Db	976	CTTTGGGATTTGCTCTCTTACTTCTACGCGAGTTTAAATCCGGCAAGAAAGTTGTGAT	1035
Qy	988	AAGACTAAGGCTCAGGATACACAACTGCTCAGTTCCTCAGTTCCTCAGTTTACTGATTATG	1047
Db	1036	AAACCTCAACCCCGGAGACAGAACTGCTATTTCTCCAGTTAGTATGTTGCTTAA	1095
RESULT 5			
AF078531		1227 bp	mRNA linear PLN 20-APR-2000
LOCUS	AF078531		
DEFINITION	Arabidopsis thaliana purine permease (PUP1) mRNA, complete cds.		
ACCESSION	AF078531		
VERSION	AF078531.1	GI:7620006	
KEYWORDS			
SOURCE	Arabidopsis thaliana (thale cress)		
ORGANISM	Arabidopsis thaliana		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.		
AUTHORS	Gillissen, B., Buerkle, L., Andre, B., Kuhn, C., Rentsch, D., Brandl, B. and Frommer, W.B.		
TITLE	A new family of high-affinity transporters for adenine, cytosine, and purine derivatives in Arabidopsis		
JOURNAL	Plant Cell 12 (2), 291-300 (2000)		
MEDLINE	20129770		
PUBMED	10662864		
REFERENCE	2 (bases 1 to 1227)		
AUTHORS	Gillissen, B., Andre, B., Rentsch, D., Buerkle, L., Kuehn, C., Brandl, B. and Frommer, W.B.		
TITLE	Direct Submission		
JOURNAL	Submitted (15-JUL-1998) Institut fuer Botanik, University of Tuebingen, Auf der Morgenstelle 1, Tuebingen, Baden-Wuerttemberg 72076, Germany		
FEATURES	Location/Qualifiers		
source	1..1227		
	/organism="Arabidopsis thaliana"		
	/mol_type="mRNA"		
	/cultivar="Columbia"		
	/db_xref="taxon:3702"		
	1..1227		
	/gene="PUP1"		
	27..1097		
gene			
CDS			



/translation="MKNGLIILINCIILITIGTCGGPLLRLLYFTNGGRIWFMFSFLSTA  
GFPILILPLVPSLRRNRNNAENRKTFLMTEPLFIASIVIGLLTCLDNL  
YSYGLAIPVSTSLIQTOLAFNALFALLVKRTPFSINAVLLTGTGILHLS  
DGDPAKESKKEVVVGLMTVAALLYAFILPLVELTYKKAQELTFPLVLEIOMVC  
LAATFCVIGMFVDPKVIAREAREFKIGGSVYVALIVITGIWQGFLEAIGIVF  
CASLASGVLSVILLPVEFAVVCVFRKFPQAEKGVSLLSLWGFVSYFYGFBSKQK  
VWDKPPPETELPILPVSXYA"

ORIGIN

Query Match	50.2%;	Score	526.4;	DB	8;	Length	1250;
Best Local Similarity	70.5%;	Pred. No.	6.8e-147;				
Matches	761;	Conservative	0;	Mismatches	286;	Indels	33;
Gaps	3;						
QY	1	AAGATGAAGATGAAGACAGTCTTGTAAATCATAAACTGTATATTTCTTTGGCCATTTGGAAC	60				
Db	44	AAGAAGATGAAGAATGGTTTGATAATCATAAACTGTATATCTCTACTATATAGGAACA	103				
QY	61	TGTGGAGCCCTCTAATGATGCGTCTCTACTTCCAAAATGGTGGCGAAGAGATCTGGTTT	120				
Db	104	TGTGGAGGTCTTTTGTAACTCTCTACTTCCACCAATGGCGGAAAACGAATCTGGTTC	163				
QY	121	CCAAGCTTCTTCAAAACCGTTGGTGTCCACTCATTTTCTTCCCTCTCTCTATCTTATCTTC	180				
Db	164	ATGAGCTTCTATCAACCGTGGTTTTTCCAATCATCTCATCTCTCTCTTGGTCTCTCTTC	223				
QY	181	CTCCGCGCTCGTCTGGTTG-----CCTTGAAGAAACAAGAAACGACTCCA	222				
Db	224	CTCAGCGTTCGCGCAGCAACCGCAACCTTAACAACGCGGAAACAAGCGGAAACAAG	283				
QY	223	TTTTTCTCATGAAACCTCTCTCTTTATCGCGCTATCGTTGTGTGTTTGTCTGTGGGA	282				
Db	284	CTCTTCTCATGAAACTCTCTGTTTATCGCCTCCATTTGTATAGGTTTGTCTCACAGGA	343				
QY	283	TTTGACAATTTACCTCTACTCTTACCGGTGAGCTTATATCCCTGTTTCTACTCGCTCTTG	342				
Db	344	CTTGACAATCTACTTATATTTCTTACGGATTAGCATATCTGCCAGTTTCAACTTCTATCGCTC	403				
QY	343	ATCATCTCCGCGCAATTAGCTTCTACGTCTCTTTGCAATTTTATATGTTGAAGCAAAAG	402				
Db	404	ATAATCGGAACCTCACTAGCTTTCACGCTCTCTTCTGCTTCTTGTGTAGTCAAGCAAAAG	463				
QY	403	TTCAACCTTTCTACTATAAACGCTATCGTTTGTCTCACTGGTGGTGGCGTAGTCTTTGCC	462				
Db	464	TTCACTCCGTTCTCCATAAACCGCGTCTGTTTGTGTACGGTTGGTATCGGGATCTTTGCG	523				
QY	463	CTTAACTCTGATAGTACAGCTTGCAGGCTTGAACGAGACACACAGGAATATGTTTGGGTTTC	522				
Db	524	TTACACAGTATGGAGACAAACCGGCTTAAGGAGAGCAAGAGATATGTTGTTGGGTTTC	583				
QY	523	ATCATGACTCTTGGTGCAGCTCTTCTCTATGGGTTTATATTTGCCACCTTGTGCGAGCTTCT	582				
Db	584	TTGATGACTGTGTTGCGAGCTCTTCTCTATGCTTTTATATATACCGCTCGTTGAGCTAACT	643				
QY	583	TACAAGAAATCTGGTCAGCGAATACAGTATACGCTCGCGCTCGAGTTCAGATGGTCTTA	642				
Db	644	TACAAGAAAGCTCGTCAAGAAATACACTTTCCCACTTGTGCTTGAGATTCAGATGTCATG	703				
QY	643	TGCTTTTGTCTGCACTTGTGTCTGCTCGTGGGAGTGTCTAGCCGCTGGGATTTCAAGGTG	702				
Db	704	TGCTTTGCTGCTACTTTTTTTCTGTGTCTATTTGGCATGTTTCATCGTTGGAGATTTTAAAGTG	763				
QY	703	ATAGCAGGGAAGCAAGAGATTTTAAAGCTTTGGAGAGTC---TTTGTACTATGTTGGTGATT	759				
Db	764	ATAGCAAGAGAGCAAGAGATTTCAAGATTTGGAGATCAGTGTTTTACTATGCAATGATA	823				
QY	760	GTGTTTCAAGCCCAATATCTGGCAAGCAATTTTGTGGGAGCTATTTGGTGTGATCTTCTGT	819				
Db	824	GTGATCACAGGAATAATATGCAAGGTTTCTTCTTAGGAGCCATAGGATTTGTTTGT	883				
QY	820	GCATCTCTCTGCTCTCTGAAATTATGGTCAGTGTCTCTGCTTCCGGTGAACGGTGATCTTG	879				
Db	884	GCAATCATAGTCTCTGCTGTTTCTGATTAAGTGTCTCTGCTTCCGGTGTACTGAAGTTTTC	943				

QY	880	GCGTCATTGCTTCCAGGAGAGTTTTCAGCGGGGAGAGTGTCTGCTTGGCTCTCTCC	939
Db	944	GCGTCGTTTGTTCGGGAGAGTTTTCAGGAGAGAGTGTCTCTCTACTTCTTCT	1003
QY	940	CTCTGGGATCAGTCTCTTATTTCTATGGACAGGTTAAATCC-----GAGGAG	987
Db	1004	CTTCTGGGATTTGTCTCTTACTTCTACGCGGAGTTTAAATCCGCGGAGAGTGTGTAT	1063
QY	988	AAGACTTAAGGCTCAGGATACACAACTGTCTCAGTCTTCCAGTTTACTGATTATGTAGCTTAA	1047
Db	1064	AAACCTCAACCGCGGAGACAGACTGCTATCTTCTGTTAGTATGTTAGTCTTAA	1123

BT005504 1102 bp mRNA linear PLN 15-MAR-2003  
Arabidopsis thaliana clone U50269 putative purine permease  
(At1g28230) mRNA, complete cds.

BT005504  
BT005504.1 GI:28973198  
FLI CDNA.  
Arabidopsis thaliana (thale cress)

Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi

Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Hsuan, V.W., Lee, J.M.,  
Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wong, C., Wu, H.C.,  
Yu, G., Yuan, S., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y.,  
Ishida, J., Jones, T., Kamiya, A., Kawai, J., Kim, C.J., Narusaka, M.,  
Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinn, P.,  
Southwick, A., Tripp, M.G., Wu, T., Shinozaki, K., Davis, R.W.,  
Ecker, J.R. and Theologis, A.

Arabidopsis Open Reading Frame (ORF) Clones  
Unpublished  
2 (bases 1 to 1102)

Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Hsuan, V.W., Lee, J.M.,  
Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wong, C., Wu, H.C.,  
Yu, G., Yuan, S., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y.,  
Ishida, J., Jones, T., Kamiya, A., Kawai, J., Kim, C.J., Narusaka, M.,  
Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinn, P.,  
Southwick, A., Tripp, M.G., Wu, T., Shinozaki, K., Davis, R.W.,  
Ecker, J.R. and Theologis, A.

Direct Submission  
Submitted (15-MAR-2003) Plant Gene Expression Center, 800 Buchanan  
Street, Albany, CA 94710, USA

The RIKEN Genomic Sciences Center (GSC) members carried out the  
collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN  
Arabidopsis Full-Length cDNA'): Seki, M., Narusaka, M., Ishida, J.,  
Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,  
Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PGEC (SSP) Consortium members constructed and  
sequenced the pUNI (ORF) clones using the RAFL cDNAs: Yamada, K.,  
Chan, M.M., Chang, C.H., Dale, J.M., Hsuan, V.W., Lee, J.M.,  
Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wong, C., Wu, H.C.,  
Yu, G., Yuan, S., Chen, H., Cheuk, R., Jones, T., Kim, C.J., Nguyen, M.,  
Palm, C.J., Shinn, P., Southwick, A., Tripp, M.G., Wu, T., Davis, R.W.,  
Ecker, J.R. and Theologis, A.

Yamada, K. (SSP/PGEC) and Seki, M. (RIKEN GSC) contributed equally  
to this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP  
/PGEC) contributed equally to this work as PIs.

Annotation based on July 2002 version of the Arabidopsis genome  
submitted to Genbank.

Location/Qualifiers  
1. 1102  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
/chromosome="1"

Source  
1. 1102  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
/chromosome="1"



Db 241 GTCTCTTATATCGCGCTGTTATTTGTAGGCATTTCTCAGGCTTTGTAACTACTCTGTATG 300  
QY 302 CTTACGGGTAGCTTATATCCCTGTTTCTACTCGTCTTTTGATCATCTCGGCSAATTAG 361  
Db 301 CATATGGTAGCTTATCTTCAGTCTTTCACAGTCTCTTATCATTTGCTTCACAGTTAG 360  
QY 362 GCTTCACTGCTCTCTTTGCACTTTTATGGTGAAGCAAAAGTTTACACACCTTTTCACTATAA 421  
Db 361 CTTTATAGCTATCTCTCTCATTTCTCATGTTTAAACATAAAGTTTCACTCTTTTACCATCA 420  
QY 422 ACCTATCTTTTGTCTCACTGGTGGTCCGTAGTCTTGTCCCTTAATCTCTGATAGTGACA 481  
Db 421 ATGCTGTGTGTTGTGACTGTTGGTCTGCGGTTTTGGGAATGSCATACGAACTGATA 480  
QY 482 AGCTTGCAACAGGAGACACACAAGGAATATGTTGTGGTTCATCATGACTCTTTGGTGCAG 541  
Db 481 AGCAGTTCTATGAGACTCACACAGTACATAACTGTTCTTGTATCTGTAGCAGCAG 540  
QY 542 CTCTTCTATGCGTTTATATTTGCCACTTGTGAGGCTTTTACAAAGAAATCTGGTCAAG 601  
Db 541 CTGTTATGATGCTTTTCACTTCCCATTTAGTGAACCTTGTACCAAGAAAGCTTAACCAA 600  
QY 602 GAATCAGTATACGCTCGCGCTCGAGTTCAGATGCTTATATGCTTTGTGCTGCCACTTGTG 661  
Db 601 CCATGAGCTATACCTTGTGCTCGAGTTCAGTGTGATTTGTGCTCTCTCTTCTTATTTG 660  
QY 662 TCTGCTCTGTTGGGATGCTAGCGCTGGGATTTCA--AGTGATAGCAGGAGAGCAA 718  
Db 661 TCAGCTCATCGTATGTTTCATCGCTGGTGAATTTCAAGCAGGCTTACCAAAAGAGCAA 720  
QY 719 GAGATTTAAGCTTGGAGAGTCTTTGTACTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 778  
Db 721 GAGAGTTCAAGCTTGGAGAGGCACTGTTCTATGTTGTTGTTGTTGTTGTTGTTGTTGTT 780  
QY 779 GGCAAGCATTTTGTGGAGCATTTGGTGTGATCTTCTGTGATCGTCTCTGTTCTGTCTGTG 838  
Db 781 GGCAAGGCTTCTTCTGGAGGCAATGGATTAATCTTCTCCACATCGTCTCTGCTCGG 840  
QY 839 GAATATATGTCAGTCTCTGCTTCCGTTGACGCTGATCTTGGCGTCATTTGCTTCTCCAGG 898  
Db 841 GTATATGATATCAGTCTTTTCCCAATACAGAGGTTTGTAGCTGTATATCTACCATG 900  
QY 899 AGAAGTTTCAGGCGGGAAGGTGTCGCTTTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 958  
Db 901 AAAAGTTTCAAGCTGAGAGGACTTCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 960  
QY 959 ATTTCTATGACAGGTTAAATCCGAGGAGAACTTAAG 996  
Db 961 ACTTTTATGTCAGATAAAGTCTGGCGAGGATAAAAGG 998

## RESULT 9

AC021044 121720 bp DNA linear PLN 30-OCT-2002  
LOCUS Arabidopsis thaliana chromosome I BAC F3H9 genomic sequence,  
DEFINITION complete sequence.

ACCESSION AC021044 GI:8347959

VERSION AC021044.5

KEYWORDS HTG.

SOURCE Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.

1 (bases 1 to 121720)

Federspiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F.,  
Altafi,H., Nguyen,M., Lam,B., Southwick,A., Mitanda,M., Brooks,S.,  
Buehler,E., Chao,O., Chin,J., Choi,E., Gonzalez,A.,  
Hwang,B., Johnson-Hopson,C., Khan,S., Kim,C., Koo,T., Lee,J.M.,  
Lenz,C., Liu,A., Liu,S., Mukharshy,N., Pham,P., Sakano,H.,  
Shinn,P., Toriumi,M., Vaysberg,M., Walker,M., Yu,G., Ecker,J.,  
and Davis,R.W.

REFERENCE  
AUTHORSJOURNAL  
REFERENCE  
AUTHORS

Unpublished  
2 (bases 1 to 121720)  
Federspiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F.,  
Altafi,H., Nguyen,M., Lam,B., Southwick,A., Bei,Q., Buehler,E.,  
Chin,C., Chiu,J., Choi,E., Dunn,P., Gonzalez,A., Hwang,B., Kim,C.,  
Koo,T., Lee,J.M., Lenz,C., Li,J., Liu,A., Liu,K., Liu,S.,  
Mukharshy,N., Pham,P., Sakano,H., Schwartz,J., Shinn,P.,  
Thaveri,A., Toriumi,M., Vaysberg,M., Walker,M., Yu,G., Ecker,J.,  
Theologis,A. and Davis,R.W.  
Direct Submission  
Submitted (12-JAN-2000) DNA Sequencing and Technology Center,  
Stanford University, 855 California Avenue, Palo Alto, CA 94304,  
USA

TITLE  
JOURNALREFERENCE  
AUTHORS

3 (bases 1 to 121720)  
Federspiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F.,  
Altafi,H., Nguyen,M., Lam,B., Southwick,A., Bei,Q., Buehler,E.,  
Chin,C., Chiu,J., Choi,E., Dunn,P., Gonzalez,A., Hwang,B., Kim,C.,  
Koo,T., Lee,J.M., Lenz,C., Li,J., Liu,A., Liu,K., Liu,S.,  
Mukharshy,N., Pham,P., Sakano,H., Schwartz,J., Shinn,P.,  
Thaveri,A., Toriumi,M., Vaysberg,M., Walker,M., Yu,G., Ecker,J.,  
Theologis,A. and Davis,R.W.  
Direct Submission  
Submitted (29-MAR-2000) DNA Sequencing and Technology Center,  
Stanford University, 855 California Avenue, Palo Alto, CA 94304,  
USA

TITLE  
JOURNALREFERENCE  
AUTHORS

4 (bases 1 to 121720)  
Federspiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F.,  
Altafi,H., Nguyen,M., Lam,B., Southwick,A., Bei,Q., Buehler,E.,  
Chin,C., Chiu,J., Choi,E., Dunn,P., Gonzalez,A., Hwang,B., Kim,C.,  
Koo,T., Lee,J.M., Lenz,C., Li,J., Liu,A., Liu,K., Liu,S.,  
Mukharshy,N., Pham,P., Sakano,H., Schwartz,J., Shinn,P.,  
Thaveri,A., Toriumi,M., Vaysberg,M., Walker,M., Yu,G., Ecker,J.,  
Theologis,A. and Davis,R.W.  
Direct Submission  
Submitted (08-JUN-2000) DNA Sequencing and Technology Center,  
Stanford University, 855 California Avenue, Palo Alto, CA 94304,  
USA

TITLE  
JOURNALREFERENCE  
AUTHORS

5 (bases 1 to 121720)  
Federspiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F.,  
Altafi,H., Nguyen,M., Lam,B., Southwick,A., Ecker,J., Theologis,A.  
and Davis,R.W.  
Direct Submission

TITLE  
JOURNAL

Submitted (11-AUG-2000) DNA Sequencing and Technology Center,  
Stanford University, 855 California Avenue, Palo Alto, CA 94304,  
USA

REFERENCE  
AUTHORS

6 (bases 1 to 121720)  
Federspiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F.,  
Altafi,H., Nguyen,M., Lam,B., Southwick,A., Ecker,J., Theologis,A.  
and Davis,R.W.  
Direct Submission

TITLE  
JOURNAL

Submitted (15-AUG-2000) DNA Sequencing and Technology Center,  
Stanford University, 855 California Avenue, Palo Alto, CA 94304,  
USA

## COMMENT

On Jun 8, 2000 this sequence version replaced gi:7340332.  
Bases 1-7,373 of clone F3H9 overlap with bases 96,792-104,163 of  
IGF BAC clone F3M18. gb|AC010155  
e-mail for correspondence: arabesequence.stanford.edu Genes with  
similarity to proteins in the databases are named 'putative',  
'-like' or 'similar to'. Genes that have EST similarity but no  
significant protein similarity are described as 'unknown proteins'.  
Genes that are annotated based only on gene prediction software  
are described as 'hypothetical proteins'. The gene prediction  
programs used to predict genes include: Grail (Informatics Group,  
Oak Ridge National Laboratory,  
http://combio.ornl.gov/section/index.html), GENSCAN (Chris Burge,  
http://genome.stanford.edu/~chris/GENSCAN.html), and NetPlantGene  
(S.M. Hebsgaard, et al., CBS, Technical University of  
Denmark, http://www.cbs.dtu.dk/NetPlantGene.html).

FEATURES  
source

1..121720  
/organism="Arabidopsis thaliana"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:3702"



```
/chromosome="1"
/misc_feature
1..7373
/note="overlap with bases 96,792-104,163 of IGF BAC clone
F3M18, gb|AC010155, see GenBank record for clone F3M18 for
annotation in this region."
gene
9852..10816
/gene="F3H9.1"
/note="synonym: DRM1"
join(9852..9957,10049..10274,10780..10816)
CDS
/gene="F3H9.1"
/note="Location of cDNA clone gb|AF053746 and genomic
clone gb|AF053747."
/codon_start=1
/product="dormancy-associated protein"
/protein_id="AAF98422.1"
/db_xref="GI:9795604"
/translation="MVLLEKLDWDVAVGPQPDRLGLRKKITQPINIRIDIGSSSK
VVMRSLTMPAAVSPGTPTTPTTPRKDNVMSVFNPGNSLATRAIGSNI PDKPTH
VNSPSVDWLYSGDSRSQHR"
complement(12229..14900)
gene
/gene="F3H9.2"
complement(join(12229..12924,13048..13089,13397..13665,
14024..14447,14511..14900))
CDS
/gene="F3H9.2"
/note="Hypothetical protein"
/codon_start=1
/protein_id="AAF98443.1"
/db_xref="GI:9795625"
/translation="MDSRDMLACEDEIYSIKNETSLVDLPFGIKPIGLKWFKLIK
RNSVDINKYKARLVAVGVYQRTGDFVEVPVPACTEKLIVDLEASNDWEIOLDVK
TAFUHLGKFTVYTRDEDFVEKEEGKLAQILOEPKESKPSVYRTLSGEFLL
VAVYDDLFTVTWTTIIDEFGEMTSNFMNSLKGTYYLIGIEVFORKEIRLNQSR
YEMILKEAWASCSLRESVVLAKSNEKEIDATGRKKVGCRLRYLLHLLSE
VTGPKSTVRIIRINKSAISLTKNPVPFHGEKGIHRRYHFTIRVENEQIEVERHGI
BQRADITKALVRMFKEMRDLIGQDELILSEIQNIQOQASVYMQTLKROAKVDS
KGMSTNIVRRDPTGOCIPRPIPKGMRKVRKHVDDDEGMSGGSKVKVRLTSSSG
DGDLSVPFENETTRIRNEKDLMLIEIFGPSTDDSDSGTEPESFMTNLSLWLM
GKRKRSNRKWTWRHILENHVGHQKTYQDVRKCLYCETLVHSFSLIGVVSMMQMD
IQVLLSMDASVIFNLQNPQVSCPKPNTLKSCEVCLILT"
20721..24395
/gene="F3H9.3"
join(20721..20817,21062..21222,21318..21365,21455..21514,
21709..21837,21988..22098,22183..22286,22430..22488,
22580..22674,22757..23426,23624..23816,23957..24048,
24148..24395)
CDS
/gene="F3H9.3"
/note="Unknown protein"
/codon_start=1
/protein_id="AAF98423.1"
/db_xref="GI:9795605"
/translation="MDVSKVSVFGRNFVAVLVKVGPSVPLSLNPNSSGILLPRDIFL
SGVAAKVLFPAGQDMALVAVSVVFPFLTLGHTSSISIQDPVKLI PGAMIEMVE
GQLEAKEAPWPAQLLSLVDVPVSAALQSLIEASSGSKDSGWDICWSLVSANSG
QPSNIEHYSGLQDLPHPNANFWAKSATRMALIGVPLSLGOPSNNFASSSKGGT
LVALGPSLSPVNFNSYPSGSIKKSLMATDVRLCPMGEGA PVFAKNGLHIGILIR
PLRQNSGVEIQLVWPGALITACSHLLLEEPSVEGKASQWSEVLVSXSDASIPAQ
VAIBKMSVCLITVNDGVWASGIIINLHGHLLINAHLEPWRKYKGSGVYGEKFPV
LAGEFSTSGKFWKQSQTLPKRAPRHHYSSVGENIREYKHNFLQTCRHRDVRVRL
LDSWTWCIPANTVYICEKQDIALIQLEYVPGKLQPI TANFSSPLGTATHVGHGFLG
PRGLSPSICSGVAVKVVHAKRLNTQSIQOEAFFPAMLETTAAVHPGSGAVLNS
SGHMLGLVSNRAGHGTVI PHLNFSPICAVLAPIFKFAEDMQNTTILQTDOPSEEL
SVIALMPSLSFKTEQSLPNLPLKLLKDGNNKQTKGSQFAKFIATQDMFVKPTKLSRD
VIPSKL"
31599..32534
/gene="F3H9.4"
31599..32534
/gene="F3H9.4"
/note="Hypothetical protein"
/codon_start=1
/protein_id="AAF98424.1"
/db_xref="GI:9795606"
/translation="MQSKNMIVASSHQOQQOQQPQPOPQLKPCPRCDSSNTKFCYYNN
YLSQPRHFKCKACKRYWTRGCTLRNVVPVGGSYRKNKRVKRPSTATTTTASTVSTNSS
SPNPHQISHESSNHHPLFYGLSDHSSCNVNLPMIPSPFSDSSSKTSSSGSESEFL
SSGFSSLSALGLGLPHQMSHDHINGSFINNSTNPKPFLLSGLFGSSMSSSSTLQHP
HKPMNNGDMLQSHLQTLASLQDLHVGNEDMKYKKGKLDQVIGSINGFMSSSSSL
DPNYYNNMNNASVVGAWLDPDTNNNVGSSLTSLI"
45230..48424
/gene="F3H9.5"
join(45230..45724,45882..45956,46051..46951,47136..47182,
47292..47365,47715..47984,48407..48424)
CDS
/gene="F3H9.5"
/note="Hypothetical protein"
/codon_start=1
/protein_id="AAF98425.1"
/db_xref="GI:9795607"
/translation="MDNLFPPSSNANSVOELSMDDNNNRSHPTTPTVYDHHQAPPH
FLPFFSYFVEQMAAVMNPQVYISECYPOIPTVTGSEFGLVGNPCWQERGFLOP
RMTKMARINRKAMSRNNSPNSFSELVDSKQRLMMLNLKSDKDSYQDS
TFONKLRVCEKELKNSDVGLGRIVLPKRDAEANLPKLSKDEGI VVQWRDVFMSQS
WSFKYFWSNKRMYVLENTGFVKQNGAEIGDPLTIYBDESKNLYFAMNGSGKQN
EGRESRERNHYEEAMLDYI PRDEBEASIAMLIGNLNDHYPIPNLMDLMTDLDLQHHQ
ATSSMPEDHAYVGSDDQVLDWL"
54225..55737
/gene="F3H9.6"
join(54225..54993,55427..55737)
CDS
/gene="F3H9.6"
/note="Unknown protein; proline rich amino acid sequence"
/codon_start=1
/protein_id="AAF98426.1"
/db_xref="GI:9795608"
/translation="MGFIGKSVLSVALMCFSTSSVFTBENVHKTOTPSLAPAPAPYH
HGHHHPVPPHHHPHPHPHPAKSPVKAPVSPVPAKPPVKPVVPTKAPVVP
KPPVPPVPPVPAKPPVPPVPTKAPVPTKAPVPTKAPVPTKAPVPTKAPVPTKAPV
KPVYPTKAPVPTKAPVPPVSPVPAKPPVPPVPTKAPVPTKAPVPTKAPVPTKAPVPT
VPPVPPKPSLVAVRGTVYCKSKYAAFNLLGAKPLEGATVKLVCKSKNNTIARETIT
DKNGYFLLPLAKVTIVNFGPRGCRYLVKSKDYKCSKVSKVFGDVGAEKPEKLKGS
Query Match
Best Local Similarity 63.78; Pred. No. 5.9e-111;
Matches 682; Conservative 0; Mismatches 313; Indels 76; Gaps 1;
Qy 2 AGATGAAGTGAAGACAGTCTTGTGTAATCAATAACTGTATATCTTGGCCATTGGAAACT 61
Db 83626 AGACAAGATGGTGAAGGCTCTTGTGATCATAAACTGCATAATCTAGCCATAGAACT 83685
Qy 62 GTGAGGCGCTCTAATCATCGGCTCTTACTTCCAAATGTGCGGAAGGATCTGGTTTC 121
Db 83686 GTGAGGCTCTTGAITATGCGTCTTACTTCAACAATGCGGGAAGGATTTGTTCT 83745
Qy 122 CAAGCTTCCTTCAACCGTTGGTTGTCCACTCATTTTCTCCCTCTTCTCTTATCTTCC 181
Db 83746 CTAGCTTCTTGAACCTGAGGCTTCTCTGTATCTTATCTCTCTGCTCTTCTTACA 83805
Qy 182 TCCGCCGTCGTGCTTGTGAAGAACAAAGAAAGACTTCCATTTTCTCATGAAACCTC 241
Db 83806 TTACCCGCGCAGAAGCAACAATGTGGGTGATAGTACAAGTTCTCTTCTTATCAACCGC 83865
Qy 242 CTCTCTTATCGCGCTTATCGTTGTTGCTGCGGATTTGACAAATACCTCTACT 301
Db 83866 GTCTCTTATCGCGCTGTTATTGTGAGGCATCTCTCAGGGTTTGAATACTTGTATG 83925
Qy 302 CTTACGGGTTAGCTTATATCCCTGTTTCTACTCGGTCTTTGATCATCTCCGCGCAATTAG 361
Db 83926 CATATGTATAGCTTATCTCCAGTTTCTACAGCTGCTCTTATCATTTGCTTCAGTTAG 83985
Qy 362 GCTTCACTGCTCTCTTTGCAATTTTATGTTGAAGCAAAAAGTTTCAACCTTTCTACTATA 421
Db 83986 CTTTTATAGCTATCTTCTCATTTCTTATGGTTAAACATAAGTTCACTCTCTTTTACCATCA 84045
Qy 422 ACGCTATCGTTTGTCTCACTGGTGGCGGTAGTCTTCCCTTAACTCTGATAGTGACA 481
Db 84046 ATGCTGTGTGTGTGACTGTGTGGTCTCGGGTTTGGGAATGATACCGAAACTGATA 84105
Qy 482 AGCTTGCAACGAGACACACAAGGAATATGTTGTTGGGTTTCATCATGACTCTTGGTGAG 541
```



Db	84106	AGGCAGTTATGAGACTCAACAGCTACATAACTGGTTCTTCTGATTACTGTAGCAGCAG	84165
Qy	542	CTCTTCTCTATGGTTTATATGCTGCTGAGCTTCTTCTACAGAAATCTGGTTCAGC	601
Db	84166	CTGTTATGATGTTCTCACTTCCATTAGTGAATCTGTTACCAAGAAAGCTAAGCAA	84225
Qy	602	GAATCAGTATACGCTCGGCTCGAGTTCCAGATGCTTTATGCTTTGCTGCCACTTGTG	661
Db	84226	CCATGAGCTATACCTTGTGCTCGAGTTCCAGTTGATTTGTTGCTCTCCTTCTTATTG	84285
Qy	662	TCGCTCTCTGGGATGCTAGCCGCTGGGATTTCAAGTGA	703
Db	84286	TCAGCTCATCGGTATGTTTCATCGCTGGTGATTTCAAGGTTAGTCACTCTTTTGAATCTG	84345
Qy	704	-----TA 705	
Db	84346	AAAACATACAAAGAGTTTTTAACAGAGACTTGAGCCCTTACTGTTTTTGCAGGCCCTTA	84405
Qy	706	GCAGGAGACGAAGAGATTTTAAGCTTGGAGAGTCTTTGTACTATGATGTTGTTCTTC	765
Db	84406	CCAAAAGAACGAAGAGATTTCAAGCTTGGAGAGCAATTTCTATGCTGGCTGTGTTT	84465
Qy	766	ACGGCCATATCTGCAAGCAATTTTGTGGAGCTATTGGTTGATCTTCTGTGCAATCG	825
Db	84466	TCAGCCATCATATGCGCAAGGCTTCTTCTGGGAGCCATTTGGATTAATCTTCCACATCG	84525
Qy	826	TCTCTGCTCTCTGGAATTATGGTCAGTGTCTGCTTCCGCTGACGGTGAATCTTGGCCGTC	885
Db	84526	TCTCTGCTCTGGGTATTAATGATATCATGCTGCTTTTCCCAATTACAGAGTTTTAGCTGT	84585
Qy	886	ATTGCTTCCAGGAGAAAGTTTCAGCGGGGAAAGGTGTGCGTTTGGCTCTCTCCCTCTCG	945
Db	84586	ATATTCTACCATGAAAGTTTCAAGCTGAGAGGGGACTTCTCTGCTCTCTCCCTTGG	84645
Qy	946	GGATCAGTCTCTATTCTATGACAGAGTTTAAATCCGAGAGAGCAAGCTAAG	996
Db	84646	GGCTTTGCTCTTACTTTTATGCTGAGATAAAGTCTGCGGAGGATAAAAGG	84696
RESULT 10			
AK066610			
LOCUS			
DEFINITION	Oryza sativa (japonica cultivar-group) cDNA clone:J013073M15, full insert sequence.		
ACCESSION	AK066610		
VERSION	AK066610.1 GI:32976628		
KEYWORDS	FLI_CDNA; CAP trapper.		
SOURCE	Oryza sativa (japonica cultivar-group)		
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.		
REFERENCE	1		
AUTHORS	The Rice Full-length cDNA Consortium, National Institute of Agricultural Sciences Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yamada, H., Hotta, I., Kojima, K., Itoh, M., Kawai, J., Ishikawa, M., Yamada, H., Ooka, H., Ohta, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuka, K., Shishiki, T. and Yamamoto, M.		
FEATURES			
FEATURES	source		
FEATURES	1..1505		
FEATURES	/organism="Oryza sativa (japonica cultivar-group)"		
FEATURES	/mol_type="mRNA"		
FEATURES	/cultivar="Nipponbare"		
FEATURES	/db_xref="taxon:39947"		
FEATURES	/clone="J013073M15"		
ORIGIN			
ORIGIN	Query Match 24.0%; Score 251.6; DB 8; Length 1505;		
ORIGIN	Best Local Similarity 55.6%; Pred. No. 4.5e-64; Indels 18; Gaps 2;		
ORIGIN	Matches 535; Conservative 0; Mismatches 409;		
Qy	20 TTCTTGTATCATAACTGTATATTCTTGGCCATTGGAACTGTGGAGGCCCTCTTAATGA 79		

TITLE

Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice

JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS

Science 301 (5631), 376-379 (2003)

22752273

12869764

2 (bases 1 to 1505)

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K., Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Hotta, I., Iida, J., Itoh, M., Iwata, R., Imanura, K., Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I., Kanagawa, S., Kato, H., Kawagashira, N., Kawai, J., Kawamata, M., Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koyama, S., Kurahara, C., Kurokawa, T., Kusumegi, T., Li, C., Lu, M., Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A., Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M., Namiki, T., Narikawa, R., Niikura, J., Nishi, K., Nomura, K., Numasaki, R., Ohneda, E., Ohno, M., Ohtsuka, K., Oka, M., Ooka, H., Osato, N., Ota, Y., Otsu, Y., Otsu, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K., Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and Yoshimura, A.

Direct Submission

Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of Agricultural Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kamondai, Teikyo, Ibaraki 305-8602, Japan [E-mail: skikuchi@nias.affrc.go.jp, Tel: 81-29-838-7007, Fax: 81-29-838-7007]

This clone is one of the 28K full-length cDNA clones from japonica rice.

URL : <http://cdna01.dna.affrc.go.jp/cDNA/>

NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuka, K., Shishiki, T. and Yamamoto, M.

FAIS Genome Sequencing & Analysis Group: Otsu, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kurokawa, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Narikawa, R., Niikura, J., Oka, M., Ry, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K. and Murakami, K.

Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Iida, J., Imanura, K., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Kanagawa, S., Kato, H., Kawai, J., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koyama, S., Kurahara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Osato, N., Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K., Yasunishi, A. and Hayashizaki, Y.

Location/Qualifiers

1..1505

/organism="Oryza sativa (japonica cultivar-group)"

/mol\_type="mRNA"

/cultivar="Nipponbare"

/db\_xref="taxon:39947"

/clone="J013073M15"

FEATURES

source

ORIGIN

Query Match 24.0%; Score 251.6; DB 8; Length 1505;

Best Local Similarity 55.6%; Pred. No. 4.5e-64; Indels 18; Gaps 2;

Matches 535; Conservative 0; Mismatches 409;

Qy 20 TTCTTGTATCATAACTGTATATTCTTGGCCATTGGAACTGTGGAGGCCCTCTTAATGA 79



Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y., Itoh, M., Kigawa, I., Kanagawa, S., Katoh, H., Kawai, J., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Oota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toyota, T., Waki, K., Yasunishi, A. and Hayashizaki, Y.

## FEATURES

Location/Qualifiers  
1..1458  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="mRNA"  
/cultivar="Nipponbare"  
/db\_xref="taxon:39947"  
/clone="001-036-F06"

## ORIGIN

Query Match 23.1%; Score 242; DB 8; Length 1458;  
Best Local Similarity 55.0%; Pred. No. 3.5e-61;  
Matches 529; Conservative 0; Mismatches 415; Indels 18; Gaps 2;  
QY 20 TTCTTGTGTAATCAATAAAGTATATTTCTTGGCCATTCGAAACTGTGGAGGCCCTCTAATCA 79  
DB 202 TCCTCTGGGGCTCAACTGGCGGATGCTGTGTGGGACACAGCGGGCGCCGATCCCTA 261  
QY 80 TGGCTCTCTACTTCCAAATATGGGGGAAAGGATCTGGTTTCCAAAGTCTCTTCAACACG 139  
DB 262 GCGGGCTCTACTTCAGCAGGGGGCCACAGGAAGTGGCTCTGGCGTGGCTCGAGACG 321  
QY 140 TTGGTTGTCACTATTTCTCCCTCTCTCTTTATCTTCTTCCTCCCGGCTGTGCTGGC 199  
DB 322 CCGGGTGGCGGCTGTGCTGTGGTGGCGGTGCTGCTTACCACAGCGCGCGCGCGG 381  
QY 200 TTGAAGAACAGAAAGCAGTCCATTTTCTCATGAAACCTCTCTTTATCGCGCTA 259  
DB 382 ACCGCGCG-----CGCCGCTGTCTCTACGCGCGGCGGCTGTGCTGGCGGGG 432  
QY 260 TCGTTGTGGTTGTGCTGGGATTTGACAAATPACCTCTACTCTTACGGGTTAGCTTATA 319  
DB 433 CGGTGCTCGGGCTCTACCGGCGCGAGACTTCGTCTACGCTACGGCTACGGCTACCG 492  
QY 320 TCCCTGTTCTACTGCTCTTGTATCATCTCGCGCAATPAGGCTTCACTGCTCTCTTTG 379  
DB 493 TGGCGGTGTCCACTCGGCGATCCTCATCTCCACGAGCTCGGGTTCAACCGCTTCTTCG 552  
QY 380 CATTTTATTTATGGTGAAGCAAAAGTTACACCTTTTCACTATAACGCTATCGTTTGTCTCA 439  
DB 553 CGTGCTCATGTCGGCGAGCGGCTCACCGCGGAGCGTCAACCGCGGTGGCGCTGCTCA 612  
QY 440 CTGGTGGTGGCTGAGTCTTGGCTTAACTCTGATAGTGAAGCTTGAACAGACAC 499  
DB 613 CTATCGCGCGCTGCTGGGCTCCACGCGTCCAGGACCGACCGCGGGGTGACCA 672  
QY 500 ACAGGAATATGTTGTTGGTTTCAATGACTCTTGGTGAGAGCTTCTTCTCTATGGGTTTA 559  
DB 673 CCGGAAGTACTGGAATGGGCTTCTTCTCACCTCGCGCGCGCGGCTGACGGGCTCA 732  
QY 560 TATTGCCACTTGTGAGCTTTCTTACAAGAAATCT-----GFTACGAGCAATCACGT 610  
DB 733 TCCTGCGGCTGCTGAGCTGCGGTACAAGACCGCGCGGGCGGCGCGGCTCACGT 792  
QY 611 ATACGCTCGCGCTCGAGTTCAGATGCTTTATGCTTTTGTGCTGCACTTGTGCTGCTCG 670  
DB 793 ACGCGCTGTGATGGAGATGACAGTGTCTATGGCTTCTTTCGCCACCGGCTCTGTACCG 852  
QY 671 TGGGGATGTAGCCGCTGGCGATTTCAAGGTGATACGAGGAAGCAAGAGATTTTAAAGC 730

DB 853 TCGGTGTGATCGTCAACAGGAGCTTCCAGCGCATCCGAGGGAGGCGAGCAGTACGAGC 912  
QY 731 TTGGAGAGCTTTTGTACTATGTGGTGATTTGTTCAAGGCCATAATCTGCAAGCAATTTT 790  
DB 913 TGGGTGAACGGGGTACTACGTGGTGGTGTTCACCGCGCTCTGTGGAGTCTTCTTCT 972  
QY 791 TTGTGGAGCTATTGCGGTTGATCTTCTGTGCATCGTCTCTGTGCTCTGGAATTTATGTC 850  
DB 973 TCGTGGCGCGCTCGGCGTCACTTCTTGGGTGCACACGCTGCTCGCGGGATCATCATCG 1032  
QY 851 GTGCTCTGCTTCGGTGAAGGATCTGGCGGCTATTTGCTTCCAGGAGAGATTTTCAGG 910  
DB 1033 CCGTGTTCATCCCGATCACGAGGCTGTGGGGTCACTTCTCTCCACGAGAAATTCAGCA 1092  
QY 911 CGGGGAAGGCTGCTCGCTTTGGCTCTCTCCCTCTGCGGATCAGTCTCTTATTTCTATGGAC 970  
DB 1093 GCGAGAAGGCGTGGCGTCTGCTCTCTGCGGCTCGCTCTGCGGCTCTCTACTCTCTACGGC 1152  
QY 971 AG 972  
DB 1153 AG 1154  
RESULT 12  
AK0641135  
LOCUS  
DEFINITION  
Oryza sativa (japonica cultivar-group) cDNA clone:002-102-G04, full insert sequence.  
ACCESSION  
AK0641135  
VERSION  
AK0641135.1 GI:32974153  
KEYWORDS  
FLI CDNA; oligo-capping.  
SOURCE  
Oryza sativa (japonica cultivar-group)  
ORGANISM  
Oryza sativa (japonica cultivar-group)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.  
REFERENCE  
1  
AUTHORS  
The Rice Full-Length cDNA Consortium, National Institute of Agricultural Sciences Rice Full-Length cDNA Project Team; Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yasaki, J., Ishikawa, M., Yamada, H., Ooka, K., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group; Otsu, Y., Murakami, K., Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y., Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M., Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niihara, J., Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J., Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN; Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Osato, N., Ota, Y., Saito, K., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Yoshino, M. and Hayashizaki, Y.  
japonica rice  
Collection, mapping, and annotation of over 28,000 cDNA clones from Science 301 (5631), 376-379 (2003)  
2752273  
PUBMED  
12869764  
TITLE  
japonica rice  
JOURNAL  
Science 301 (5631), 376-379 (2003)  
MEDLINE  
2752273  
PUBMED  
12869764  
REFERENCE  
2 (bases 1 to 1363)  
AUTHORS  
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K., Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imamura, K., Kagawa, I., Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kawai, J., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M., Masuda, H., Murakami, K., Matsuyama, T., Miura, J., Miyazaki, A., Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M., Namiki, T., Narikawa, R., Niihara, J., Nishi, K., Nomura, K., Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H.,

Osato, N., Ota, Y., Otomo, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and Yoshimura, A.

Direct Submission  
Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan [E-mail: skikuchi@nias.affrc.go.jp, Tel. 81-29-838-7007, Fax: 81-29-838-7007]

This clone is one of the 28K full-length cDNA clones from japonica rice.  
URL : <http://cdna01.dna.affrc.go.jp/cDNA/>  
NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Sato, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T., and Yamamoto, M.

FAIS Genome Sequencing & Analysis Group: Otomo, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Narikawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K., and Murakami, K.

Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Oso, N., Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K., Yasunishi, A. and Hayashizaki, Y.

TITLE  
JOURNAL

COMMENT

FEATURES  
source

ORIGIN

Query Match 21.5%; Score 226; DB 8; Length 1363;  
Best Local Similarity 53.7%; Pred. No. 2.3e-56;  
Matches 536; Conservative 0; Mismatches 430; Indels 32; Gaps 2;  
  
QY 9 GATGAAGACAGTCTTGTAATCATAAACGTGTATATCTTGGCCATTGGAACTGTGGAGG 68  
DB 173 GTTCAGAGCCGCTCTCGTGTGCTCACTTCTGCTCATGTGTGCTCGGCTGGCGG 232  
  
QY 69 CCCTCTAATGATGCGTCTCTACTTCCAAAATGGTGGCGAAAGGATCTGGTTTCCAAGCTT 128  
DB 233 GCGGCTCTCTCCGCGCTACTTCTCCGCGGGGCAACCGCAAGTGGCTCTCCAGGCT 292  
  
QY 129 CTTCAAAACCGTGTGTGTCACCTATTTCTTCCCTCTTCTATATCTTCTCCCGCG 188  
DB 293 TCTCAGACCCGCGGCTGCGGCTGCTCTGCGCGGCTCTGCTTCTCGTACTCTCTCAGC 352  
  
QY 189 TCGTGTGCTTGTGAAGAACCAAGAA-----CGACTCCATT 224  
DB 353 CGCGCGCGCGGAGGTGTGAGACGACGGCGCTGGCGCGGCGGCGGCGGCGCT 412  
  
QY 225 TTTCCTCATGAACCTCTCTCTTTTATCGCCGCTATCGTTGTTGCTCGTGGGATT 284  
|||||

DB 413 GTTCTCATGACGCCACGCTCTCGTGGCGTCCGCGCTCGTGGGCTCATGACCGCGGT 472  
QY 285 TGACAAATTACTCTTACTCTTAAGGTTAGCTTATATCCCTGTTTCTACTGCTTTCGAT 344  
DB 473 CGACGACCTCTCTACGCTACGCGCTGGCTACTCTCCGGTGTCCACCTCTCCATCCT 532  
QY 345 CATCTCCGCGCAATTAGCTTTCAGTCTCTCTTTGCAATTTTATGTTGAAGCAAAAGTT 404  
DB 533 CATCTCCAGGAGCTGGCTTTCACGGCGGCTTCCGCTGCTGCTCGTGCCACGCGTT 592  
QY 405 CACACCTTTCAATAAAGCTATCGTTTGTCTCACTGGTGGTGGTGGTGGTGGTGGTGGT 464  
DB 593 CACGCGCTTCTCGTGAACGCGCTGCTGCTCAGCGTCCGCGCGGCGATGCTGGGAT 652  
QY 465 TAACCTGTATGATGACAAAGCTTGAACGAGACACAAAGGAATATGTTGTTGGTTCAT 524  
DB 653 GAACCGCGCGGAGACCGCGCGGGGCTGTCGCGCGCGAGTACTGCGCGCGGTTGCG 712  
QY 525 CATGACTTGTGTCAGCTCTTCTATAGGTTTATATTTGCC-----ACTTGTGAG 576  
DB 713 CATGACGCTGCGCGCGGGGCTGTAGGGCTGCTGCTCCCTCATGGAGCTCAGCAG 772  
QY 577 CTTTCTTCAAGAAATCTGTCAGCGAATCACGATACGCTCGCGCTCGAGTTCCAGATG 636  
DB 773 GCGCACCGCGCGCGCGCGCGCTCAGGTACAGCTCGTCTCATGGAGATGACGCTC 832  
QY 637 GTCTTATGCTTGTGTCGCACTTGTGCTGCTGCTGGGATGCTAGCGGCTGGCGATTTTC 696  
DB 833 GTCATCGGCTTCTGTCGCGCGCGCTTTCAGCGCGCTGCGCATGCTCGTCAACACGATTTTC 892  
QY 697 AAGGTGATAGCAGGAGAAAGAGATTTTAAAGCTTGGAGAGTCTTTGTACTATGTGGTG 756  
DB 893 CACCAATCCAGAGAGAGCCCATGTTCCGCTGCGGCAAGCGGCTACTACTGCTC 952  
QY 757 ATTGTGTTCAAGCCATTAATCTGGCAAGCAATTTTGTGGAGCTATTTGGGTTGATCTTC 816  
DB 953 CTGCGCGGATCGGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1012  
QY 817 TGTCATCGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 876  
DB 1013 TACGGCTCGGCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1072  
QY 877 TTGCGCGCTCATTTGCTTCCAGGAGAGTTCAGCGCGGGAAGGTGTCGCTTTTGGCTCTC 936  
DB 1073 CTGCGCGCTCATGTTCTTCCAGAGCAATTCAGCGCAAGAGGCGCTGCGCTGCTGCTC 1132  
QY 937 TCCCTCTGGGATCAGTCTCTTATTTCTATGGACAGGT 974  
DB 1133 TCGCTATGGGCTTCTGCTCTCTACTTCTACGCGGAGGT 1170

RESULT 13  
AY096558  
LOCUS  
DEFINITION Arabidopsis thaliana putative protein (At4g18200) mRNA, complete cds.  
ACCESSION AY096558  
VERSION AY096558.1  
KEYWORDS GI:20465496  
SOURCE Fuj. CDNA.  
ORGANISM Arabidopsis thaliana (thale cress)  
Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
1 (bases 1 to 1165)  
Yamada, K., Ban, J., Chan, M.M., Chang, C.H., Chang, E., Dale, J.M., Deng, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Bowser, L., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Shinn, P., Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.

TITLE  
JOURNAL  
REFERENCE  
AUTHORS

Arabidopsis Open Reading Frame (ORF) Clones

Unpublished  
2 (bases 1 to 1165)  
Yamada,K., Banh,J., Chan,M.M., Chang,C.H., Chang,E., Dale,J.M.,  
Deng,J.M., Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L.,  
Tang,C.C., Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L.,  
Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J.,  
Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B.,  
Lin,J., Meyers,M.C., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J.,  
Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A.,  
Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.  
Direct Submission  
Submitted (16-APR-2002) Plant Gene Expression Center, 800 Buchanan  
Street, Albany, CA 94710, USA  
The RIKEN Genomic Sciences Center (GSC) members carried out the  
collection and clustering of RAPL cDNAs (RAPL cDNA : 'RIKEN  
Arabidopsis Full-Length cDNA'): Seki,M., Narusaka,M., Ishida,J.,  
Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai, J.,  
Hayashizaki,Y. and Shinozaki,K.

TITLE  
JOURNAL  
COMMENT

The Salk, Stanford, PGENC (SSP) Consortium members constructed and  
sequenced the pUNI (ORF) clones using the RAPL cDNAs: Yamada,K.,  
Banh,J., Chan,M.M., Chang,C.H., Chang,E., Dale,J.M., Deng,J.M.,  
Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C.,  
Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L., Chen,H.,  
Cheuk,R., Jones,T., Karlin-Neumann,G., Kim,C., Lam,B., Lin,J.,  
Meyers,M.C., Miranda,M., Nguyen,M., Palm,C.J., Shinn,P.,  
Southwick,A., Davis,R.W., Ecker,J.R. and Theologis,A.

Yamada,K. (SSP/PGENC) and Seki, M. (RIKEN GSC) contributed equally  
to this work. Shinozaki,K. (RIKEN GSC) and Theologis,A. (SSP  
/PGENC) contributed equally to this work as PIs.

Annotation is based on the January 2002 version of the Arabidopsis  
genome submitted to GenBank.

FEATURES  
source

Location/Qualifiers  
1..1165  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
/chromosome="4"  
/clone="U18528"  
/ecotype="Columbia"  
/notes="This clone is in pUNI 51."  
1..1165  
/gene="At4g18200"  
1..1134  
/genes="At4g18200"  
/codon\_start=1  
/evidence=experimental  
/product="putative protein"  
/protein\_id="AA20208.1"  
/db\_xref="GI:20465497"  
/translation="MGISQVHYCNGDQNLNLENDHEETSFSPQTKNCKRLRVSI  
YAFVFCPLATVLYENGKSTVYVLLQLGFPVLILPFRFRIROKPSDT  
NFSQSPFTLASVILCTGLLVSAVAYLSAVGLLPLVFSLLIASQLAFAPFSYF  
LNSOKPTFLLVNSLPLLVNTSDENTNVSQVYVIGFTCTIGASAGIGLV  
LSLIOLFRKVTKHTSSAVLDLANQSLVATCVLIGLFSAGEWRTLPSEMNKLG  
KYSYLITLASAIFPQVYVGVGLIFESSVSFNSITAVGLPIPVVAVIVFDKMD  
ASKIFSIILAIWGLFVSFYQHYLDLDEKKLTKCTQKPEEETQL"  
1135..1165  
/gene="At4g18200"

## 3' UTR

## ORIGIN

Query Match 16.0%; Score 167.6; DB 8; Length 1165;

Best Local Similarity 48.6%; Pred. No. 9e-39;

Matches 492; Conservative 0; Mismatches 514; Indels 6; Gaps 1;

QY 1 AGAGTGAAGATGAAGACAGTCTCTGTAATCATAACTGTATATCTTGGCCATTGGAAAC 60  
|||||  
Db 100 AGAACTGTAAAGGTGGCTCGGTGCTCATATACGCAATCTTGTCTATCTTCTGCCAA 159  
QY 61 TGTGGAGGCCCTCTAATGATGCGTCTCTACTTCCAAAATGGTGGCGAAAGGATCTGGTTT 120

Db 160 CCATTGCTACAGTCTCTGGTAGACTGATATGAAATGGAGGAAACACATATGTG 219  
QY 121 CCAAGCTTCCCTTCAAAACGTTGGTTGTCACACTCATTTTCTTCCCTCTCTCTTATCTTTC 180  
Db 220 GTAACACTTCTTCAACTCATTTGGCTTCCCTGTAAGTATCTCTCCGCTTCTTTTCTCGA 279  
QY 181 CTCGCCGCTGCTGCTTGGCTTTGAAGAAACAAGAAACAGACTCCATTTTTTCTCATGAACCT 240  
Db 280 ATCAGGCAAGCCAAATCAACAGATACAAATTTTCAGTCAGTCCCTTCCCTTCAACCCCTT 339  
QY 241 CCTCTCTTTATCGCCGCTATCGTTGTTGTTGCTGCTGGGATTTGACAAATTTACCTCTAC 300  
Db 340 GCATCGGTTTAC-----TTGTGCACCTGGACTGCTAGTCTCTGCTTATGCTTTATTTGCT 393  
QY 301 TCTTACGGGTTAGCTTATATCCCTGTTCTACTGCTCTTTGATCATCTCTCGCGGAATTA 360  
Db 394 GAGTTGGTGGCTTACTTACCAGTCTCTACTTTCTCCCTCATCTTGGCCTCACAGTTG 453  
QY 361 GGCCTTCACTGCTCTCTTTGCAATTTTTTATGTTGTAAGAAAGTTTCACACCTTTTCACTATA 420  
Db 454 GCCTTCACTGCTCTTTTCTCATATTTTCTTAACTCGCAAAAGTTTCACCTCTTGTAGTGC 513  
QY 421 AAGCTATCGTTTTCCTCACTGCTGGTGGCGTAGTCTTTCCTTAACTCTGATAGTAGAC 480  
Db 514 AATTCTTTGTTTCTCTTACGGTTTCTCTGCTGCTCTCTGCTGCTCAACACTGATTCAGAA 573  
QY 481 AAGCTTGCAGAACGACACACAAAGAAATATGTTGTTGGTTCATCATGACTCTTGTGTGCA 540  
Db 574 AACACAAATATGATCTAGAGTACAAATGATGATGGGTTTCACTGTACCAATTTGTGTCT 633  
QY 541 GCTCTTCTCTATGGGTTTATATTTGCCACTTGTGCAGCTTTCTTCAAGAAATTCGTGTACG 600  
Db 634 TCCGCTGGGATTTGGATTGTAATCTCTGATACAACTGCTCTTCAGGAAAGTTTTCACG 693  
QY 601 CGAATCAGTATACCTCGGCTCGAGTTCCAGATGCTTATGCTTGTGCTGCCACTTGT 660  
Db 694 AAGCATACATCTCAGCAGTCTCTGCACTTGGCCAAATTACCAGTCTCTAGTTGCAACTTGT 753  
QY 661 GTCTGCTCGTGGGGATGCTAGCCGCTGCGCATTTTCAAGGTGATAGCAGGAGAACAAGA 720  
Db 754 GTGTACTACTAGGACTGTTTGCAAGTGGAGTGGAGAACTCTGCCAAGTAGATGAGA 813  
QY 721 GATTTTAAAGCTTGGAGAGTCTTTGTACTATGTGTGTGATTTGTGTTTCCAGGCCATAATCTGG 780  
Db 814 AACTACAACTGGGGAAGTGTCTATATCTTGACTTTGGCCTCAGCAGCTATTTTCTGG 873  
QY 781 CAAGCATTTTGTGGGAGCTATTGGGTTGATCTTCTGTGCAATGCTCTCTGTGCTCTGGA 840  
Db 874 CAAGTATACACTGTTGTTGTGGGATTAATCTTCGAGTCTTCTTCTGTGTTCTCCAAT 933  
QY 841 ATTATGCTCAGTCTCTGCTTCCGGTGACGGTGAATCTTGGCGCTCATTTGCTTCCAGGAG 900  
Db 934 TCCATAACTGCTGTGGGAGTGCCTATATGTTCCAGTTGTAGCAGTATAGTTTTTCCATGAT 993  
QY 901 AAGTTTCAAGCGGGGAAAAGTGTCCCTTTGGCTCTCTCCCTCTGGGGATCAGTCTCTTAT 960  
Db 994 AAGATGGATGCATCAAGATTTTCTCCATCATTTTAGCTATCTGGGGCTTCTCTTCAATTT 1053  
QY 961 TTCTATGACAGGTTTAAATCCGAGGAGAAGACTAAGGCTCAGGATACACAAC 1012  
Db 1054 GTCTATCAGCACTACCTCGAGCAAAAGATTGAAGACTTGGCAGACAAAAC 1105

## RESULT 14

AY074546

LOCUS

Arabidopsis thaliana 1428 bp mRNA linear PLN 18-SEP-2002

DEFINITION

cds.

ACCESSION

AY074546

VERSION

AY074546.1

KEYWORDS

FLI CDNA.

SOURCE

Arabidopsis thaliana (thale cress)

ORGANISM	Arabidopsis thaliana		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.		
AUTHORS	1 (bases 1 to 1428)		
	Yamada,K., Liu,S.X., Sakano,H., Pham,P.K., Banh,J., Chung,M.K., Goldsmith,A.D., Lee,J.M., Quach,H.L., Toriumi,M., Yu,G., Bowser,L., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B., Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.		
TITLE	Arabidopsis Full Length cDNA Clones		
REFERENCE	2 (bases 1 to 1428)		
AUTHORS	Yamada,K., Banh,J., Chan,M.M., Chang,C.H., Chang,E., Dale,J.M., Deng,J.M., Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B., Lin,J., Meyers,M.C., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.		
TITLE	Direct Submission		
JOURNAL	Submitted (16-JAN-2002) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA		
COMMENT	RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAPL cDNAs (RAPL cDNA : 'RIKEN Arabidopsis Full-Length cDNA': Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K.		
	The Salk, Stanford, PGEC (SSP) Consortium members carried out the sequencing and annotation of the RAPL cDNAs: Yamada,K., Banh,J., Chan,M.M., Chang,C.H., Chang,E., Dale,J.M., Deng,J.M., Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L., Cheuk,R., Jones,T., Karlin-Neumann,G., Kim,C., Lam,B., Lin,J., Meyers,M.C., Miranda,M., Nguyen,M., Palm,C.J., Shinn,P., Southwick,A., Davis,R.W., Ecker,J.R. and Theologis,A.		
	Yamada,K. (SSP/PGEC) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Theologis,A. (SSP/PGEC) contributed equally to this work as PIs.		
FEATURES	Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.		
source	Location/Qualifiers		
	1..1428		
	/organism="Arabidopsis thaliana"		
	/mol_type="mRNA"		
	/db_xref="taxon:3702"		
	/chromosome="4"		
	/clone="RAF109-36-N15 (R18628)"		
	/ecotype="Columbia"		
	/note="This clone is in a modified pBluescript vector (FLC-1) as a BamHI/XhoI insert."		
gene	1..1428		
	/gene="At4g18200"		
5'UTR	1..71		
	/gene="At4g18200"		
CDS	72..1205		
	/gene="At4g18200"		
	/codon_start=1		
	/evidence=experimental		
	/product="unknown protein"		
	/protein_id="AA169512.1"		
	/db_xref="GI:1849121"		
	/translation="MGISQVHCYNGDONLEANDHRETESFSVPOTKNCKRLRVSI YAIIVFCOPLATVILGRVYENGKSTVYVTLLOLIGPVLILFRFSRIRQPKSTDT NFSGSPSTTLASVYLCTGLVSAVYLSAVGLYLPVSTFSLIASQLAFPTFFSF LNSQKFTPLVNSLFLTVSALLIVNDSTNTNVRQVIGFICTIGASAGILG LSLIQLFRKVFTHKTSVAVLDLANYSQSLVATCVLLIGLFASGEWRTLPSEMRNYKLG		
	KVSYILTLASAAAFQWYTVGCVGLIFESSVSFNSITAVGLPIVPVAVIVFHKMD ASKIFSIILAIWGLSPVYQHYLDEKXKLTQTKPVEEETQL"		
	1206..1428		
	/gene="At4g18200"		
	/misc_difference 1412		
	/gene="At4g18200"		
	/note="not present in genomic sequence"		
ORIGIN			
Query Match	16.0%; Score 167.6; DB 8; Length 1428;		
Best Local Similarity	48.6%; Pred. No. 9.3e-39;		
Matches 492; Conservative 0; Mismatches 514; Indels 6; Gaps 1;			
QY	1	AAGATGAAGATGAAGACAGATTCTTGTAAATCATAACTGTATATTCTTGGCCATTGGAAAC	60
DB	171	AAGAACTGTGAAGAGTGGCTCCGTGTCCATATACGCAATCTTTGTCACTCTTCGCCAA	230
QY	61	TGTGGAGGCCCTTAATGATGCGTCTCTACTTCCAAATGGTGGCGAAAGATCTGGTTT	120
DB	231	CCACTTGTACAGTTCTGGGTAGACTGTACTATGAAATGGAGGAAAGCACATATGTG	290
QY	121	CCAAGCTTCCTTAAACCGTTGGTGTGCACCTCAATTTTCTCCCTCTTCTCTTATCTTC	180
DB	291	GTAACACTTCTTCAACTCAITGGCTTCCCTGTACTGATTCTCTCCGCTTCTTTCTCGA	350
QY	181	CTCCGCGTGTGCTTGCCTTGAAGAAACAAGAACGACTCCATTTTCTCATGAACCT	240
DB	351	ATCAGGCAACCAATCAACAGATACAAATTCAGTCAGTCCCCTTCTTCCACACCCCT	410
QY	241	CCTCTTTATTCGCGCTATGTTTGTGGTGTGCTCGTGGGATTTGCAATTAACCTCTAC	300
DB	411	GCATCGGTTTAC-----TTGTGCACGGAGCTGCTGTCTTATGCTTATTTGTCT	464
QY	301	TCITACGGGTAGCTTATATCCCTGTTTCTACTCGCTTGTGATCATCTCGCGCAATTA	360
DB	465	GCAGTTGGGTGTCTTACTTACCAGTCTCTACTTTCTCCCTCATCTTGGGCTCACAGTTG	524
QY	361	GGCTTCACTGCTCTCTTTGCAATTTTATGGTGAAGCAAAAGTTTCAACCTTTTCACTATA	420
DB	525	GCCTTCACTGCTTTTCTCATATTCTTAACTCGCAAAAGTTCACTCTCTTGTATGTC	584
QY	421	AACGCTATCGTTTGTCTACTGGTGGTGGCTAGTCTCTGCCCTTAACTCTGATAGTAC	480
DB	585	AATCTTTTGTCTTCTTACGGTTTCTCTGCCCTCTCTCGTGGTCAACACTGATTCAGAA	644
QY	481	AAGCTTCCAACGAGACACACAGGATATATGTTTGGGTTTCATCATGACTCTTGGTGCA	540
DB	645	AACACAAACAAATGTATCTAGAGTACAATATGTGATTGGGTTCATCTGTACCATTGGTCT	704
QY	541	GCTCTTCTATGGTTTATATTGCCACTTGTGCGAGCTTTCTTACAAGAAATCTGGTCAG	600
DB	705	TCCGCTGGATTGGATTGGTACTATCTCTGATACAACTGCTCTTCAGAAAGTTTTCAG	764
QY	601	CGAATACGTATACGCTCGCGCTCGAGTTCAGATGGTCTTATGCTTTTGTGCGCACTGT	660
DB	765	AAGCATACATCCTCAGCAGTCTGAGACTTGGCCAAATACCACTCTCTAGTTGCAACTGT	824
QY	661	GTCTGCCCTCGTGGGGATGCTAGCGCTGGCGATTTCAGGTGATAGCAGGAGAGCAGA	720
DB	825	GTGGTACTCATAGGACTGTTTGAAGTGGAGAGTGGAGAACTCTGCCAAGTGAGATGAGA	884
QY	721	GATTTTAAAGCTTGGAGAGTCTTTGTACTATGTGTGATTGTGTTCAGGGCCATAATCTGG	780
DB	885	AACCTACAACTGGGGAAGTGTCAATATCTTGAATTTGGCTCAGCAGCATTTTCTGG	944
QY	781	CAAGCATTTTGTGGAGCTATTGGGTTCATCTTCTGTGCACTGCTCTGGTCTCTCGGA	840
DB	945	CAAGTATACACTGTTGGTGTGGGATTAATCTTCAGTCTTCTTCTGTGTGTTCTCCAAT	1004
QY	841	ATTATGGTCACTGCTGCTTCCGTTCCGTTACCGGTGATCTTGGCCCTCATTTGCTTCCAG	900
DB	1005	TCCATAACTGCTGTGGGACTGCCTATAGTTTCCAGTTGTAGCAGTGATAGTTTTCATGAT	1064



**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 1, 2004, 02:44:43 ; Search time 4791.13 Seconds  
(without alignments)  
10669.751 Million cell updates/sec

Title: US-09-913-767-6  
Perfect score: 1081  
Sequence: 1 caaatccacagttcaagat.....ataagtgatattttgaaag 1081

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues  
Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*

- 1: gb\_ba.\*
- 2: gb\_htg.\*
- 3: gb\_in.\*
- 4: gb\_om.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sts.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1077.8	99.7	1081	6	AX033549 Sequence
2	1043.8	96.6	1047	6	AX412868 Sequence
3	1008.8	93.3	82697	8	ATT9A21
4	1008.8	93.3	194143	8	AL021713 Arabidops
5	974.8	90.2	1106	8	AF370622 Arabidops
6	835.8	77.3	1071	6	AX033550 Sequence
7	406.8	37.6	1293	6	AX033547 Sequence
8	406.8	37.6	58427	8	ATF15U5
9	406.2	37.6	3387	6	AX507517 Sequence
10	406.2	37.6	3387	6	AX52015 Sequence
11	385.2	35.6	73179	8	AP006383 Lotus cor
12	369.6	34.2	1194	6	AX033548 Sequence
13	368.2	34.1	1165	8	AY096558 Arabidops
14	368.2	34.1	1428	8	AY074546 Arabidops
15	307	28.4	1476	8	AK071254 Oryza sat
16	307	28.4	1617	8	AK072660 Oryza sat
17	307	28.4	175645	8	AP004096 Arabidops
18	298.8	27.6	1251	8	BT003358 Arabidops
19	298.8	27.6	1372	8	AY062782 Arabidops

20	298.8	27.6	115421	8	T12C22	AC020576 Sequence
21	298.4	27.6	1145	6	AX033546 Sequence	AX033546 Sequence
22	294	27.2	110036	8	AC142095	AC142095 Medicago
23	262	24.2	133299	8	OSJN00098	AL606658 Oryza sat
24	262	24.2	145739	8	OSJN00157	AL662957 Oryza sat
25	258.8	23.9	1486	8	AK119483	AK119483 Oryza sat
26	234.2	21.7	83698	6	AX540653	AX540653 Sequence
27	234.2	21.7	83698	6	AB010072	AB010072 Arabidops
28	225.8	20.9	1117	8	AY096468	AY096468 Arabidops
29	225.8	20.9	1333	8	AY074535	AY074535 Arabidops
30	225.8	20.9	1403	8	AY087844	AY087844 Arabidops
31	225.8	20.9	104386	8	AT132A17	AL161813 Arabidops
32	225.8	20.9	199362	8	ATCHRIV24	AL161512 Arabidops
33	222.8	20.6	2175	6	AX654201	AX654201 Sequence
34	211.6	19.6	1152	6	AX653666	AX653666 Sequence
35	208.2	19.3	721	11	BX664674	BX664674 Arabidops
36	205	19.0	1279	8	BT014359	BT014359 Lycopersi
37	186.2	17.2	733	8	AF083706	AF083706 Arabidops
38	180.4	16.7	1361	8	AK071680	AK071680 Oryza sat
39	145.6	13.5	1049	6	AX033545	AX033545 Sequence
40	138.2	12.8	1067	6	AX033553	AX033553 Sequence
41	136	12.6	1225	6	AX033544	AX033544 Sequence
42	136	12.6	1227	8	AF078531	AF078531 Arabidops
43	134.4	12.4	1102	8	BT005504	BT005504 Arabidops
44	134.4	12.4	1250	8	AK117664	AK117664 Arabidops
45	129.2	12.0	1155	6	AX653139	AX653139 Sequence

ALIGNMENTS

RESULT 1	AX033549	Sequence 6 from Patent WO049152.	1081 bp	DNA	linear	PAT 21-SEP-2000
LOCUS	AX033549	Sequence 6 from Patent WO049152.	1081 bp	DNA	linear	PAT 21-SEP-2000
DEFINITION	AX033549	Sequence 6 from Patent WO049152.	1081 bp	DNA	linear	PAT 21-SEP-2000
ACCESSION	AX033549	Sequence 6 from Patent WO049152.	1081 bp	DNA	linear	PAT 21-SEP-2000
VERSION	AX033549.1	GI:10280293	1081 bp	DNA	linear	PAT 21-SEP-2000
KEYWORDS	Arabidopsis thaliana (thale cress)					
ORGANISM	Arabidopsis thaliana					
REFERENCE	Andre B., Buerkle, L., Frommer, W.B. and Gillissen, B.					
AUTHORS	Nucleic acids that code for a nucleobase transporter					
TITLE	Patent: WO 049152-A 6 24-AUG-2000;					
JOURNAL	ANDRE BRUNO (BE) ; BUERKLE LUKAS (DE) ; FROMMER WOLF B (DE) ; GILLISSEN BERND (DE)					
FEATURES	Location/Qualifiers					
source	1..1081					
	/organism="Arabidopsis thaliana"					
	/mol_type="unassigned DNA"					
	/db_xref="taxon:3702"					
ORIGIN	Query Match					
	Best Local Similarity 99.7%; Score 1077.8; DB 6; Length 1081;					
	Matches 1079; Conservative 0; Mismatches 25; Indels 0; Gaps 0;					
Qy	1 CAAATCCACAGTTCAAGATGAAAGAAATTCAGTCAGTAGCAGCGCAAGCAAGTATCTC 60					
Db	1 CAAATCCACAGTTCAAGATGAAAGAAATTCAGTCAGTAGCAGCGCAAGCAAGTATCTC 60					
Qy	61 ACTCTAACACATACAAACGGTGGCTCAGGAGTACTATGACACGGAGGAAACAGTAAA 120					
Db	61 ACTCTAACACATACAAACGGTGGCTCAGGAGTACTATGACACGGAGGAAACAGTAAA 120					
Qy	121 TGGCTAGCAACGGTAGTTCAACTTGTGGCTTTCTCTGCTACTTCCATATATATCTTG 180					
Db	121 TGGCTAGCAACGGTAGTTCAACTTGTGGCTTTCTCTGCTACTTCCATATATATCTTG 180					
Qy	181 TCATTTAAACACATGCAACAACTGATAGATGAAAAAGAAACCTCACCTAGGAACCGT 240					

|||||  
181 TCATTTAAACACATGCAACCACTGATAGATAGATGGAAGAAAGAACCTCACCTAGAGACCGT 240  
QY  
241 GTATTGGTTTACGTAGTCTTGGACTTCTTGTAGAGCAGATTGCTATCTCTACTCCATT 300  
Db  
241 GTATTGGTTTACGTAGTCTTGGACTTCTTGTAGAGCAGATTGCTATCTCTACTCCATT 300  
QY  
301 GGACTTCTTTACTTACCCGTTTCTACCTATTCCCTGATCTGTGCAATCTCAGTTAGCCCTTC 360  
Db  
301 GGACTTCTTTACTTACCCGTTTCTACCTATTCCCTGATCTGTGCAATCTCAGTTAGCCCTTC 360  
QY  
361 AATGCTTCTTCTCTTATTTCTTAACCTCAACAAAACCTTACCCCTATCATTTTAAATTCCT 420  
Db  
361 AATGCTTCTTCTCTTATTTCTTAACCTCAACAAAACCTTACCCCTATCATTTTAAATTCCT 420  
QY  
421 CTTTTCCTCTTAACCTATATCTTCCACCCCTACTTGCATTTCAATAATGAGGACAGACTCC 480  
Db  
421 CTTTTCCTCTTAACCTATATCTTCCACCCCTACTTGCATTTCAATAATGAGGACAGACTCC 480  
QY  
481 ACAAAAGTTACAAAGGAGATATGTCAAAGTTTTCATATGCACCGTTGCTGCTGCT 540  
Db  
481 ACAAAAGTTACAAAGGAGATATGTCAAAGTTTTCATATGCACCGTTGCTGCTGCTGCT 540  
QY  
541 GGTATGCTCTAGTCTTATCTTATCCCTACACAGCTAGCCCTTTCTAAAAGTCTCTAAAAGCAAA 600  
Db  
541 GGTATGCTCTAGTCTTATCTTATCCCTACACAGCTAGCCCTTTCTAAAAGTCTCTAAAAGCAAA 600  
QY  
601 AATTTCCTCAGAGTTATGGATATGATAATCTACGTGAGTCTAGTGGCCAGTTGTTAGC 660  
Db  
601 AATTTCCTCAGAGTTATGGATATGATAATCTACGTGAGTCTAGTGGCCAGTTGTTAGC 660  
QY  
661 GTGTGGGGCTTTTGTCTAGCAGTGAGTGGAAAACCTTGGACAGTGAATGGATAACTAC 720  
Db  
661 GTGTGGGGCTTTTGTCTAGCAGTGAGTGGAAAACCTTGGACAGTGAATGGATAACTAC 720  
QY  
721 AAAACATGGGAAGGTATCTTACATTATGAACCTAGTGTGGACAGCTGTTTACCTGCAGTTA 780  
Db  
721 AAAACATGGGAAGGTATCTTACATTATGAACCTAGTGTGGACAGCTGTTTACCTGCAGTTA 780  
QY  
781 TTCTCCATCGGTGGCAGAGGACTGATCTTCAGAGCTCTCCTCTCTATTTCTCAAAATGCAATA 840  
Db  
781 TTCTCCATCGGTGGCAGAGGACTGATCTTCAGAGCTCTCCTCTCTATTTCTCAAAATGCAATA 840  
QY  
841 AGCGTTTGGGACTCCCAAGTGGTCTCTATCTTGGCTGTAACTATTTCCATGACAAAATG 900  
Db  
841 AGCGTTTGGGACTCCCAAGTGGTCTCTATCTTGGCTGTAACTATTTCCATGACAAAATG 900  
QY  
901 AATGGTTTAAAGGTGATTTCTATGATCTTAGTATTTGGGGTTTACCTTCTATGCTCTAC 960  
Db  
901 AATGGTTTAAAGGTGATTTCTATGATCTTAGTATTTGGGGTTTACCTTCTATGCTCTAC 960  
QY  
961 CAACAATATCTTGATGACAAAACCTTGAAGAAAATCATGAAATCACAACAACAGAAATCC 1020  
Db  
961 CAACAATATCTTGATGACAAAACCTTGAAGAAAATCATGAAATCACAACAACAGAAATCC 1020  
QY  
1021 CCTGACCCACCAAGAGCAGAGTCACTTGGCAATCAAAATAGCTGATATTTTGAAA 1080  
Db  
1021 CCTGACCCACCAAGAGCAGAGTCACTTGGCAATCAAAATAGCTGATATTTTGAAA 1080  
QY  
1081 G 1081  
Db  
1081 G 1081

RESULT 2  
AX412868  
LOCUS AX412868 1047 bp DNA linear PAT 14-JUN-2002  
DEFINITION Sequence 632 from Patent WO0222675.  
ACCESSION AX412868  
VERSION AX412868.1 GI:21445326  
KEYWORDS  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
1  
Glazebrook,J., Wang,X., Dangl,J.L., Eulgem,T. and Zhu,T.  
Plant genes, the expression of which are altered by pathogen  
infection  
Patent: WO 022675-A 632 21-MAR-2002;  
Syngenta Participations AG (CH) ; UNIVERSITY OF NORTH CAROLINA AT  
CHAPEL HILL (US) ; Glazebrook, Jan (US) ; Wang, Xun (US) ; Dangl,  
Jeffrey L. (US) ; Eulgem, Thomas (US)  
FEATURES  
source 1..1047  
/organism="Arabidopsis thaliana"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:3702"  
ORIGIN  
Query Match 96.6%; Score 1043.8; DB 6; Length 1047;  
Best Local Similarity 99.8%; Pred. No. 2.8e-245;  
Matches 1045; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 19 ATGAAGAAAATTCAGTCAGTAGCAGCAAGCAGAGATATCTCACTCTAAACATACAAAC 78  
Db 1 ATGAAGAAAATTCAGTCAGTCAGTAGCAGCAAGCAGAGATATCTCACTCTAAACATACAAAC 60  
QY 79 GTGGCTCAGAGTATATGATGACGAGGAGAAACAGTAAATGGCTAGCAACGGTAGTT 138  
Db 61 GTGGCTCAGAGTATATGATGACGAGGAGAAACAGTAAATGGCTAGCAACGGTAGTT 120  
QY 139 CAATCTGTCGCTTCCCTGCTACTTCCATATTTATATCTTGCATTTTAAACACATGCA 198  
Db 121 CAATCTGTCGCTTCCCTGCTACTTCCATATTTATATCTTGCATTTTAAACACATGCA 180  
QY 199 ACAACTGATAGATGAAAGAAACCTCACCTAGGAACCGTGTATTTGGTTTACGTAGTG 258  
Db 181 ACAACTGATAGATGAAAGAAACCTCACCTAGGAACCGTGTATTTGGTTTACGTAGTG 240  
QY 259 CTTGGACTTCTGTAGAGCAGATTGCTATCTGTACTTCCATTTGGACTTCTTTACTTACCC 318  
Db 241 CTTGGACTTCTGTAGAGCAGATTGCTATCTGTACTTCCATTTGGACTTCTTTACTTACCC 300  
QY 319 GTTCTACCTATTCCTCGATCTGTGCACTCTCAGTTAGCCCTCAATGCTTTCTTCTCTTAT 378  
Db 301 GTTCTACCTATTCCTCGATCTGTGCACTCTCAGTTAGCCCTCAATGCTTTCTTCTCTTAT 360  
QY 379 TTTCTTAACCTCAAAAACCTTACCCCTATCATTTTAAATTTCTCTTTTCTCTTAACCTATA 438  
Db 361 TTTCTTAACCTCAAAAACCTTACCCCTATCATTTTAAATTTCTCTTTTCTCTTAACCTATA 420  
QY 439 TCTTCCACCTACTTGCATTTCAATATGAGGAGACAGACTCCACAAAAGTTTACAAAAGGA 498  
Db 421 TCTTCCACCTACTTGCATTTCAATATGAGGAGACAGACTCCACAAAAGTTTACAAAAGGA 480  
QY 499 GAGTATGTCAAAAGTTTTCATATGCACCGTTGCTGCGTCTGCTGTTATGTTAGTCTTATA 558  
Db 481 GAGTATGTCAAAAGTTTTCATATGCACCGTTGCTGCGTCTGCTGTTATGTTAGTCTTATA 540  
QY 559 TCCCTACAACAGCTAGCCCTTTCTAAAAGTCTCTAAAAGAGCAAAATTTCTCAGAAAGTTATG 618  
Db 541 TCCCTACAACAGCTAGCCCTTTCTAAAAGTCTCTAAAAGAGCAAAATTTCTCAGAAAGTTATG 600  
QY 619 GATATGATATCTACGTAGTCTAGTGCCAGTTGTTAGCTGTGGGGCTTTTGGCT 678  
Db 601 GATATGATATCTACGTAGTCTAGTGCCAGTTGTTAGCTGTGGGGCTTTTGGCT 660  
QY 679 AGCAGTCAGTGGAAAACCTTTTGGCAGTGAATGATAACTACAAAACATGGAAGGTATCC 738  
Db 661 AGCAGTCAGTGGAAAACCTTTTGGCAGTGAATGATAACTACAAAACATGGAAGGTATCC 720  
QY 739 TACATTATGAACCTAGTGTGGACAGCTGTTTACCTGGCAGTTATTTCTCCATCGGTGGCACA 798  
Db 721 TACATTATGAACCTAGTGTGGACAGCTGTTTACCTGGCAGTTATTTCTCCATCGGTGGCACA 780





QY 902 ATGGGTTAAAGTGATTCTATGATCCTAGCTATTGGGGTTTCACTTCTATGCTACC 961  
|||||  
Db 16226 ATGGGTTAAAGTGATTCTATGATCCTAGCTATTGGGGTTTCACTTCTATGCTACC 16285  
|||||  
QY 962 ACAAATATCTTGATGACAAAACCTTGAAGAAAATCATGAATCACAACAGATCCC 1021  
|||||  
Db 16286 ACAAATATCTTGATGACAAAACCTTGAAGAAAATCATGAATCACAACAGATCCC 16345  
|||||  
QY 1022 CTGACCCACGAGAGAGAGTCAACTTGGCAATCAAAATAAGCTGATATTTTGAAG 1081  
|||||  
Db 16346 CTGACCCACGAGAGAGAGTCAACTTGGCAATCAAAATAAGCTGATATTTTGAAG 16405  
|||||

## RESULT 4

## ATCHRIV48

LOCUS ATCHRIV48 194143 bp DNA linear PLN 16-MAR-2000  
DEFINITION Arabidopsis thaliana DNA chromosome 4, contig fragment No. 48.  
ACCESSION AL161548

VERSION AL161548.2 GI:7268604

## KEYWORDS

SOURCE Arabidopsis thaliana (thale cress)

## ORGANISM

Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

## REFERENCE

1 (bases 1 to 111084)  
Murphy,G., Ridley,P., Hudson,S., Mewes,H.W., Lemcke,K. and

## AUTHORS

Mayer,K.F.X.

## JOURNAL

Unpublished

## REFERENCE

2 (bases 107966 to 194143)  
Hilbert,H., Braun,M., Holzer,E., Brandt,A., Duesterhoeft,A.,

## AUTHORS

Mewes,H.W., Lemcke,K. and Mayer,K.F.X.

## JOURNAL

Unpublished

## REFERENCE

3 (bases 1 to 194143)  
EU Arabidopsis sequencing, project.

## AUTHORS

Direct Submission

## TITLE

Submitted (10-MAR-2000) MIPS, at the Max-Planck-Institut fuer

## JOURNAL

Biochemie, Am Klopferspitze 18a, D-82152 Martinsried, FRG, E-mail:  
lemcke@mips.biochem.mpg.de, mayer@mips.biochem.mpg.de, Project

## REFERENCE

Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge

## AUTHORS

Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK,  
E-mail: michael.bevan@bbsrc.ac.uk

## COMMENT

Information on performance of analysis and a more detailed

annotation of this entry and other sequences of chromosomes 3, 4

and 5 can be viewed at: <http://www.mips.biochem.mpg.de/proj/tha/>

this fragment has an overlap with ATCHRIV47 at the 5' end and an

overlap with ATCHRIV49 at the 3' end.

## FEATURES

## source

1. .194143  
/organism="Arabidopsis thaliana"  
/mol\_type="genomic DNA"  
/variety="Columbia"  
/db\_xref="taxon:3702"  
/chromosome="4"  
/gene="AT4g18120"  
6377..9293  
join(6377..6796,7122..7238,7347..7501,7578..8456,  
8543..8627,8716..8784,8871..9293)  
/note="similarity to Me12-like protein, Arabidopsis  
thaliana, gb:D86122  
contains EST gb:W43784, A1998063.1"  
/codon\_start=1  
/product="putative protein"  
/protein\_id="CAB78814.1"  
/db\_xref="GI:7268605"  
/db\_xref="GOA:Q9SVV9"  
/db\_xref="TrEMBL:Q9SVV9"  
/translation="WASGLDHFSGGIGNLMDGDSHPIGNMLPDDEELFSLMDLLN  
LSSLPATLDDLEDDYDLFSGGGLELETDPYSLNKGFRMGFADSNVNPQIFQN  
GVGSAGHPHYGEHPSRILFVRNINNSVDSELQALFDNFSEKDNVGTLLVFNLA  
SVSNRDLNIFGVYGEIREQIRETNPGRHKKFVEFDFVRSADAALKALNRTEIARI

KLEHSPGCAARRNMLQMNPELEQDDSYSLNHVESPLASSPIGHWNRNSPIDHPLQSF  
QOPQFGLNSPTKNIRYDFEFMKTASVNDQGRFRFSLDHLFSSSYNNASHKASTF  
QOPQFSGVSGFLGNSHPSPSHVETLSGSEFLWGPSSSSAWPNVPTSSNREHFPYSA  
QNGSLQHLHIGSAPSGFFPRSPETSSGVSAPFCASGNMNAORNLRETSSSNFKMLS  
APRSQTLFTGNSYLPAAATWVSIDDPEDSGNQFDSNGNQADIKIQFQLDSKIMR  
GEDPRTTUMIKNIPNKYTRNMLLAAIDBKNSGTDFLYLPIDFKNKSNVGYAFINNV  
PKFTIALYEAFNGKKWDFSEKSEKASLAYARIQGAALIAHFQNSLMNDRRCQPIV  
FDGSSKYFVNLLQOQTHDKMIYLFVCVILMNLTHTCLEFTRSGSFPAAASNSTG  
FLIFLSARTHIKADSQNVNHTASGLKRYQFLG"

exon

/gene="AT4g18120"

/number=1

intron

/gene="AT4g18120"

/number=1

exon

7122..7238

/gene="AT4g18120"

/number=2

intron

7239..7346

/gene="AT4g18120"

/number=2

exon

7347..7501

/gene="AT4g18120"

/number=3

intron

7502..7577

/gene="AT4g18120"

/number=3

exon

7578..8456

/gene="AT4g18120"

/number=4

intron

8457..8542

/gene="AT4g18120"

/number=4

exon

8543..8627

/gene="AT4g18120"

/number=5

intron

8628..8715

/gene="AT4g18120"

/number=5

exon

8716..8784

/gene="AT4g18120"

/number=6

intron

8785..8870

/gene="AT4g18120"

/number=6

exon

8871..9293

/gene="AT4g18120"

/number=7

gene

9594..13230

/gene="AT4g18130"

complement(join(9594..9837,9944..10237,10323..11130,

11238..13230))

/gene="AT4g18130"

complement(join(9594..9837,9944..10237,10323..11130,

11238..13230))

/gene="AT4g18130"

/note="Contains ATP/GTP-binding site motif A (P-loop)

AA634-641;Phytochrome chromophore attachment site

AA317-326

contains EST gb:T21020, AA395462"

/codon\_start=1

/product="phytochrome E"

/protein\_id="CAB78815.1"

/db\_xref="GI:7268606"

/db\_xref="GOA:P42498"

/db\_xref="Swiss-Prot:P42498"

/translation="MGFESSSAASNMKPPQKNTAQYSDAALFADFAQSIYTGKS

FNYKSVISPNHVPDEHITAYLSNIQKGLVQPGCLIAVEEPSFRILGSDNSDF

LGLLSLSTSHSGEFDKVKGLIGIDARTLFTPSSGASLSKAASFTSLNPLVHSR

TTQKPFYALLHRIDAGIVMDLEPAKSGDPALTLAGAQSKLAVRAISRLOSLEGDI

GALCDVVEDVQRUTGDRVMVYQFHEDDHGEVVEIRSDLEPLGLHGYPATDIPQA

ARFLFQKRNVRMI CDCNATPVKVVSSELKPLCLVNSTLRAPGHGCHTQYMANMGSA

SLALAIYVVKGDKSKLWGLVVGHCSPRYVFPFLRYACEFLMQAGLQLOLQASQ

LAEKAMRTOTLLICDMLLRDTVSALVTQSPGIMDLVKCDGAALYYKKGKMLVGVTPNE  
SOVKDLVNLVNHGDDSDSTGLTTDLSLVDAGYPGALSLGDAVCGVAAPERSKDYLLWF  
RSNTASAIKWGAKHHPKDDAGRMHPSRSTFAPLEVAKSRSLPWEISEIDAIHSLR  
LIMRESFTSVSGNGVARDANELTSFVCEMVRVLETATAPIFGVDDSSGCIINGWKK  
KTAMETGLLASEANGLSLADRIVOBESRAALESLLKALQGESEKSVMLKURFGQNN  
HPDYSSDVIVNSCTSDRYENIIGVCFVGQDITSBKAITDRFIRLQGDYKTTVQSL  
NPLIPIIPASDENACCSWNAAMEKLIGWSKHEVIGKMLPGEVFGVCVKVCODSLTK  
FLISLYGIADNPVPESSLVFFFNKEGYIEASLTANKSNIEKSVIRCPFFLIQINK  
ESGUSCPLEKSAQSLNLTIVROEIKNPLNGIRFAHKLLESSEISASQRFLETSDA  
CEKQITTIETSDLRKSBEEGLQLETFEFLRENLDITIIISQVMIIILRERNQSQFVEA  
BEIKTLPNGDRVKLQITLADLRNINVNHAPFPNSWVGISISPGQELSRRNGRIYIHQ  
FRMLHPGKGLSEMLSDMFETRDGWTDPDGLGLKSLRKLKLBQMNGRVSYVREDRCFF  
OVDLQVKMLGVESRGTEGSSSIK"  
complement (9594..9837)  
/gene="AT4g18130"  
/number=1  
intron  
complement (9838..9943)  
/gene="AT4g18130"  
/number=1  
exon  
complement (9944..10237)  
/gene="AT4g18130"  
/number=1  
intron  
complement (10238..10322)  
/gene="AT4g18130"  
/number=2  
exon  
complement (10323..11130)  
/gene="AT4g18130"  
/number=3  
intron  
complement (11131..11237)  
/gene="AT4g18130"  
/number=3  
exon  
complement (11238..13230)  
/gene="AT4g18130"  
/number=4  
gene  
14593..16312  
/gene="AT4g18140"  
gene  
complement (join (14593..14751,14921..15070,15338..15639,  
15915..16055,16141..16312))  
/gene="AT4g18140"  
CDS  
complement (join (14593..14751,14921..15070,15338..15639,  
15915..16055,16141..16312))  
/gene="AT4g18140"  
/note="similarity to HYA22 protein, Homo sapiens,  
PIR2:JC5707  
contains EST gb:R65235"  
/codon\_start=1  
/product="putative protein"  
/protein\_id="CAB78816.1"  
/db\_xref="GI:7268607"  
/translation="MKSLIKDSLRKRVLGCKPKHVDVKNKSCFDPKISRMELLVY  
ATQCNNEQNIGLDSSRVNEDQDLIVEDSSKNASNMDTIFSPVLDDLYSET  
DEFSVLDAEAQYMLLPYLEDRSASDDVKSDENRINRSRKNLEAAENHTEAQT  
EDFPOIELRNOPELADVFNKVPDMQOPRDSPKRKAVTFVLIIWMLDILDPDGKFS  
QRFRVDSGILSDGIYTKDLTVLGLDLAKVAIVDNCQPVYRLQINNGPIKSWYDDPTD  
DGLTLPLFLTLADNDVRPIAKRFNKE"  
complement (14593..14751)  
/gene="AT4g18140"  
/number=1  
intron  
complement (14752..14920)  
/gene="AT4g18140"  
/number=1  
exon  
complement (14921..15070)  
/gene="AT4g18140"  
/number=2  
intron  
complement (15071..15337)  
/gene="AT4g18140"  
/number=2  
exon  
complement (15338..15639)  
/gene="AT4g18140"

Query Match 93.3%; Score 1008.8; DB 8; Length 194143;  
Best Local Similarity 94.6%; Pred. No. 8e-237;

		Matches 1079;	Conservative	0;	Mismatches	2;	Indels	59;	Gaps	1;
Qy	1	CAAAATCCAA	CAGTTC	CAAGATG	AAAGAAATTC	CAGTCAGT	CAGCAGCC	AGCAGAG	AGTAGT	CTCTC
Db	43653	CAAAATCCAA	CAGTTC	CAAGATG	AAAGAAATTC	CAGTCAGT	CAGCAGCC	AGCAGAG	AGTAGT	CTCTC
Qy	61	ACTCTAAC	ACATAC	CAAAACGGT	GGCTCAG	-----	-----	-----	-----	88
Db	43713	ACTCTAAC	ACATAC	CAAAACGGT	GGCTCAGT	GGCTCAGT	GGCTCAGT	GGCTCAGT	GGCTCAGT	CTCTC
Qy	89	-----	-----	-----	-----	-----	-----	-----	-----	43772
Db	43773	GCCTAAC	CAGTTC	TACAAATTTT	GGGCAGAGT	TACTAT	GACAAACGG	AGGAAAC	CAGTAAAT	43832
Qy	122	GGCTAG	CAACGGT	AGTTC	CAACTTTT	GGCTTT	GGCTTT	GGCTTT	GGCTTT	CTCTC
Db	43833	GGCTAG	CAACGGT	AGTTC	CAACTTTT	GGCTTT	GGCTTT	GGCTTT	GGCTTT	CTCTC
Qy	182	CATTTAA	AAACACAT	GC	CAACCACT	GTAGAGAT	GG	AAAAA	GAACCT	CACTAGG
Db	43893	CATTTAA	AAACACAT	GC	CAACCACT	GTAGAGAT	GG	AAAAA	GAACCT	CACTAGG
Qy	242	TATTGG	TTTACG	TAGTCT	TGGACTT	CTCT	TCTAG	GAGCAG	ATTGCT	ATCTCT
Db	43953	TATTGG	TTTACG	TAGTCT	TGGACTT	CTCT	TCTAG	GAGCAG	ATTGCT	ATCTCT
Qy	302	GACTTCT	TACTTAC	CCGTTT	CTACCTAT	TTCCTG	ATCTCT	GATCTG	CATCTC	AGTCC
Db	44013	GACTTCT	TACTTAC	CCGTTT	CTACCTAT	TTCCTG	ATCTCT	GATCTG	CATCTC	AGTCC
Qy	362	ATGCTTT	CTCTCT	TATTTT	CTTAACT	CACAAA	AACTTAC	CCCTAT	CATTTAA	AAATCTC
Db	44073	ATGCTTT	CTCTCT	TATTTT	CTTAACT	CACAAA	AACTTAC	CCCTAT	CATTTAA	AAATCTC
Qy	422	TTTTCT	CTTAACT	ATATCT	TCCACCC	CTATCT	TCATTC	CAATAAT	AGAGAC	AGATCCA
Db	44133	TTTTCT	CTTAACT	ATATCT	TCCACCC	CTATCT	TCATTC	CAATAAT	AGAGAC	AGATCCA
Qy	482	CAAAAG	TTACAAA	AGGAGAT	TGTCT	CAAGTTT	TCATAT	GCACCG	TTGCTG	CGCTCTG
Db	44193	CAAAAG	TTACAAA	AGGAGAT	TGTCT	CAAGTTT	TCATAT	GCACCG	TTGCTG	CGCTCTG
Qy	542	GTATAG	TCTAGT	CTTAT	CCCTAC	CAACAGCT	TAGCCTT	CTTAA	AGAGTCTT	AAAGCAAA
Db	44253	GTATAG	TCTAGT	CTTAT	CCCTAC	CAACAGCT	TAGCCTT	CTTAA	AGAGTCTT	AAAGCAAA
Qy	602	ATTTCT	CAGAGT	TATG	GATATG	ATAATCT	ACG	TGAGTCT	AGTGGC	AGTGGTAT
Db	44313	ATTTCT	CAGAGT	TATG	ATAATCT	ACG	TGAGTCT	AGTGGC	AGTGGTAT	44372
Qy	662	TGGTGG	GGCTTT	TGCTAG	CAGT	GGAGAA	ACTTTG	AGCAGT	GAATG	GAATACTACA
Db	44373	TGGTGG	GGCTTT	TGCTAG	CAGT	GGAGAA	ACTTTG	AGCAGT	GAATG	GAATACTACA
Qy	722	AACATG	GGAAAG	GTAT	CCTAC	ATAATG	AACCT	TAGTGT	GGACAG	CTGTGTTACCT
Db	44433	AACATG	GGAAAG	GTAT	CCTAC	ATAATG	AACCT	TAGTGT	GGACAG	CTGTGTTACCT
Qy	782	TCCTCA	TCG	TGGC	ACAGG	ACTG	ATCTTC	GAGCTCT	CCTCTCT	CTATATTC
Db	44493	TCCTCA	TCG	TGGC	ACAGG	ACTG	ATCTTC	GAGCTCT	CCTCTCT	CTATATTC
Qy	842	GCCTTT	GGGACT	CCCGAG	TGGTTC	CTATCT	TGGCTG	TAATCAT	TTTCCAT	TTCGACAAAATGA
Db	44553	GCCTTT	GGGACT	CCCGAG	TGGTTC	CTATCT	TGGCTG	TAATCAT	TTTCCAT	TTCGACAAAATGA
Qy	902	ATGGGT	TTAAAG	GTGAT	TTTCT	TATGAT	CTCTAG	CTATTT	GGGGT	TTTCACTTCT
Db	44613	ATGGGT	TTAAAG	GTGAT	TTTCT	TATGAT	CTCTAG	CTATTT	GGGGT	TTTCACTTCT
Qy	962	AACAAT	ATCT	TGATG	ACAAAA	CTTGA	AGAAAA	ATCAT	GAAT	CACAAACAG
Db	44673	AACAAT	ATCT	TGATG	ACAAAA	CTTGA	AGAAAA	ATCAT	GAAT	CACAAACAG

QY 1022 CTGACCCACGAGCAGAGAGTCAACTTGGCAATCAAAATAGCTGATATTTGAAAG 1081  
|||||  
Db 44733 CTGACCCACGAGCAGAGAGTCAACTTGGCAATCAAAATAGCTGATATTTGAAAG 44792  
|||||

## RESULT 5

AF370622  
LOCUS Arabidopsis thaliana putative protein (T9A21.60) mRNA, complete cds.  
DEFINITION AF370622 1106 bp mRNA linear PLN 30-APR-2001

ACCESSION AF370622  
VERSION AF370622.1 GI:113877726  
KEYWORDS FLI\_CDNA.

SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

1 (bases 1 to 1106)

AUTHORS Lam,B., Southwick,A., Karlin-Neumann,G., Nguyen,M., Miranda,M.,  
Palm,C.J., Bowser,L., Jones,T., Banh,J., Chen,H., Cheuk,R.,  
Chung,M.K., Kin,C., Lin,J., Liu,S.X., Pham,P.K., Sakano,H.,  
Shinn,P., Yamada,K., Ecker,J., Theologis,A. and Davis,R.W.

TITLE Direct Submission  
JOURNAL Submitted (17-APR-2001) DNA Sequencing and Technology Center,  
Stanford University, 855 California Avenue, Palo Alto, CA 94304,  
USA

COMMENT e-mail for correspondence: arab@sequence.stanford.edu

This clone was isolated by RT-PCR.

## FEATURES

source Location/Qualifiers

1..1106  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
/chromosome="IV"  
/clone="C531"  
/ecotype="Columbia"

gene 1..1106

CDS 657..1106  
/genes="T9A21.60"  
/genes="T9A21.60"

/codon\_start=1  
/product="putative protein"  
/protein\_id="AAK43941.1"  
/db\_xref="GI:113877726"  
/translation="MDMIYVSLVASCVSVGLFASSEWKTLSSEMDNYKHGKVSIM  
NLWTATVQVFSIGTGLIFELSLFSNAISVLGLPVVPILAVIIFHDKMGLKVIS  
MLAIWGFSTVYVQYLDLDDNKKNHETTTESPPPEAEESTWQSK"

## ORIGIN

Query Match 90.2%; Score 974.8; DB 8; Length 1106;  
Best Local Similarity 99.8%; Pred. No. 2.4e-228;  
Matches 975; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 89 GAGTATACATGACACGAGGAGAAACAGTAATGGCTAGCAACGGTAGTTCAACTGTG 148  
Db 130 GAGTATACATGACACGAGGAGAAACAGTAATGGCTAGCAACGGTAGTTCAACTGTG 189  
QY 149 GCTTTCCTGTGCTTCCATATATATCTGTCATTATTAATTAACACATGCAACATGATA 208  
Db 190 GCTTTCCTGTGCTTCCATATATATCTGTCATTATTAATTAACACATGCAACATGATA 249  
QY 209 GAGTGGAAAGAACCTCACCTAGAACCGGTGATTGGTTACGTAGTCTGGACTTC 268  
Db 250 GAGTGGAAAGAACCTCACCTAGAACCGGTGATTGGTTACGTAGTCTGGACTTC 309  
QY 269 TTGTAGGACAGATTCTCTATCTGACTCCATTGGACTTCTTACTTACCCGTTTCTACCT 328  
Db 310 TTGTAGGACAGATTCTCTATCTGACTCCATTGGACTTCTTACTTACCCGTTTCTACCT 369  
QY 329 ATTCCCTGATCTGTGCATCTCAGTTAGCCCTTCAATGCTTCTCTTATTTCTTAAT 388  
|||||

Db 370 ATTCCCTGATCTGTGCATCTCAGTTAGCCCTTCAATGCTTCTCTCTTCTTCTTAATCT 429  
QY 389 CACAAAACTTACCCCTATCATTTTAAATTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCT 448  
Db 430 CACAAAACTTACCCCTATCATTTTAAATTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCT 489  
QY 449 TACTTGCATTCATTAATAGGAGAGACAGACTCCACAAAAAGTTACAAAAAGGAGAGTAGTGA 508  
Db 490 TACTTGCATTCATTAATAGGAGAGACAGACTCCACAAAAAGTTACAAAAAGGAGAGTAGTGA 549  
QY 509 AAGGTTTCATATGACCCGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 568  
Db 550 AAGGTTTCATATGACCCGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 609  
QY 569 AGCTAGCCCTTTCTAAAGTCCTAAAGAACCAAAATTTCTCAGAAGTTATGGATATGATAA 628  
Db 610 AGCTAGCCCTTTCTAAAGTCCTAAAGAACCAAAATTTCTCAGAAGTTATGGATATGATAA 669  
QY 629 TCTACGTGAGTCTAGTGGCCAGTTGTGTTAGCGTGGTGGGGCTTTTGTAGCAGTGAGT 688  
Db 670 TCTACGTGAGTCTAGTGGCCAGTTGTGTTAGCGTGGTGGGGCTTTTGTAGCAGTGAGT 729  
QY 689 GGAAAACTTTGAGCAGTGAATGGATTAATACTACAAACATGGGAAGGTATCTTACATTATGA 748  
Db 730 GGAAAACTTTGAGCAGTGAATGGATTAATACTACAAACATGGGAAGGTATCTTACATTATGA 789  
QY 749 ACCTAGTGTGGACAGCTGTTACCTGCGAGTTATCTCCATCGTGGCAGGACTGATCT 808  
Db 790 ACCTAGTGTGGACAGCTGTTACCTGCGAGTTATCTCCATCGTGGCAGGACTGATCT 849  
QY 809 TCGAGCTCTCTCTCTCTATTTCTCAAAATGCAATGAAGGGTTTGGGACTCCCGAGTGGTTCCTA 868  
Db 850 TCGAGCTCTCTCTCTCTATTTCTCAAAATGCAATGAAGGGTTTGGGACTCCCGAGTGGTTCCTA 909  
QY 869 TCTTGGCTGTAATCATATTTTCCATGACAAAATGAATGGGTTAAAGGTGATTTCTATGATCC 928  
Db 910 TCTTGGCTGTAATCATATTTTCCATGACAAAATGAATGGGTTAAAGGTGATTTCTATGATCC 969  
QY 929 TAGCTATTGGGGTTTCACTTCTCTATGCTACCAACAATATCTTATGATGACAAAACCTGA 988  
Db 970 TAGCTATTGGGGTTTCACTTCTCTATGCTACCAACAATATCTTATGATGACAAAACCTGA 1029  
QY 989 AGAAAAATCATGAAATCAACAACAGAAATCCCTGACCCACCAAGAGAGAGTCAA 1048  
Db 1030 AGAAAAATCATGAAATCAACAACAGAAATCCCTGACCCACCAAGAGAGAGTCAA 1089  
QY 1049 CTTGGCAATCAAAATAA 1065  
Db 1090 CTTGGCAATCAAAATAA 1106

## RESULT 6

AX033550

LOCUS

Sequence 7 from Patent WO049152.

DEFINITION

AX033550

AX033550.1 GI:10280294

VERSION

KEYWORDS

SOURCE

ORGANISM

Arabidopsis thaliana (thale cress)

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

1

REFERENCE

AUTHORS

TITLE

JOURNAL

Patent: WO 0049152-A 7 24-AUG-2000;

ANDRE BRUNO (BE) ; BUEKLE LUKAS (DE) ; FROMMER WOLF B (DE) ;

GILLISSEN BERND (DE)

Location/Qualifiers

1..1071

/organism="Arabidopsis thaliana"

/mol\_type="unassigned DNA"

Db	1040	GTGACCAATCAAAATAAGCTGTTACTTCAA	1070
RESULT 7			
AX033547			
LOCUS	AX033547	1293 bp	DNA linear PAT 21-SEP-2000
DEFINITION	Sequence 4 from Patent WO049152.		
ACCESSION	AX033547		
VERSION	AX033547.1	GI:10280291	
KEYWORDS			
SOURCE	Arabidopsis thaliana (thale cress)		
ORGANISM	Arabidopsis thaliana		
REFERENCE	Andre, B., Buerkle, L., Frommer, W. B. and Gillissen, B.		
AUTHORS	Nucleic acids that code for a nucleobase transporter		
TITLE	Patent: WO 0049152-A 4 24-AUG-2000;		
JOURNAL	ANDRE BRUNO (BE) ; BUEKLE LUKAS (DE) ; FROMMER WOLF B (DE) ;		
FEATURES	GILLISSEN BERND (DE)		
source	1. .1293		
ORIGIN	/organism="Arabidopsis thaliana"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:3702"		
Query Match	37.6%;	Score 406.8;	DB 6; Length 1293;
Best Local Similarity	64.8%;	Pred. No. 4.9e-89;	
Matches	622;	Conservative 0;	Mismatches 332; Indels 6; Gaps 1;
QY	89	GAGTATACATGACAAACGGAGGAAACAGTAATAATGGCTAGCAACCGGTAGTTCAACTTGTG	148
DB	293	GAGTTTACTATGAATAATGGTGGAAAGAGTACATGGATGGGAACACTTGTCCAATAATCG	352
QY	149	GCTTTTCCTGTGCTACTTCCATATATATATCTTTGTTCATTTTAAACACATGCAACAACACTGATA	208
DB	353	GCTTCCCTGTGTTCTGTGTTCTCTTCCGCTTCTTTTCCAAACCCAAACACCAACAG	412
QY	209	GAGATGAAAAAGAACCTCACCT-----AGGAACCGGTGATTTGGTTTACGTAGTGTG	262
DB	413	AAGCAGATTTTCAGAAAGTCTCTTCTTCCACCATTCTTGGATCAGTTTACATGCTTACTG	472
QY	263	GACTTCTGTAGGACAGATTGCTATCTGTACTTCCATTGCACTTCTTTACTTACCCGTTT	322
DB	473	GACTATTAGTGTCTGCTAACTCTTATATGTCCTCTGTGTGTTTACTATATACTTACCACTT	532
QY	323	CTACCTATTCCTGATCTGTGCACTCTCAGTTAGCTTCAATGCTTTCTTCTTATTTTC	382
DB	533	CTACTTTCTCCCTCATCTTGGCCTCACAATTTGGCCTTCACTGCTTCTTCTCATATTTTC	592
QY	383	TTAACTCAGAAAACTTACCCCTATCATTTTTTAAATTTCTTTTCTTCTTAACTATATCTT	442
DB	593	TAAACTCGCAGAAAGTTTCACACTTTTCATTGTGAAATTTCTGTGTTCTCTTACTATTTCT	652
QY	443	CCACCTCTTCTTCATTCAATAATGAGGAGACAGACTCCACAAAAAGTTTACAAAAAGAGAGT	502
DB	653	CTGCCCTCTCTGTGGTCAACACTGATTCGAAAAACACAGCAAAAGTGTCTAGAGTAAAT	712
QY	503	ATGTCAAAAGTTTTCATATGCAACCGTTGCT	







Db 57762 AGTTGAAGACTAGCCACACAGTCTCTGTAGGAGATCTCTACTACTCTGTGAGGAAG 57821  
|||||  
RESULT 9  
AX507517 3387 bp DNA linear PAT 27-SEP-2002  
LOCUS Sequence 2212 from Patent WO0216655.  
DEFINITION AX507517  
ACCESSION AX507517.1 GI:23388754  
VERSION  
KEYWORDS Arabidopsis thaliana (thale cress)  
SOURCE Arabidopsis thaliana  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
1  
REFERENCE  
AUTHORS Harper, J.F., Kreps, J., Wang, X. and Zhu, T.  
TITLE Stress-regulated genes of plants, transgenic plants containing  
same, and methods of use  
JOURNAL Patent: WO 0216655-A 2212 28-FEB-2002;  
The Scripps Research Institute (US); Syngenta Participations AG  
(CH)  
FEATURES  
source Location/Qualifiers  
1..3387  
/organism="Arabidopsis thaliana"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:3702"  
ORIGIN  
Query Match 37.6%; Score 406.2; DB 6; Length 3387;  
Best Local Similarity 64.7%; Pred. No. 6.5e-89;  
Matches 622; Conservative 0; Mismatches 333; Indels 6; Gaps 1;  
QY 88 GGAGTATCTATCACAACGAGGAGAAACAGTAATGCTAGCAACGGTAGTTCACACTTGT 147  
Db 226 GTAGTTTACTATGAAATGTTGGGAGAGTACATGATGGGAACACTTGTCCAACTAATC 285  
QY 148 GGCTTTCTGTGCTACTTCCATATTATCTCTCTCATTTTAAACACATGCAACACTGAT 207  
Db 286 GGCTTCCCTGTCTGTTCTCTCCGCTTCTTTCCCAACCAAAATCCCAACCAACA 345  
QY 208 AGAGATGGAAAGAACCTCACCT-----AGAACCGGTGATGTGTTACGTAGTGGT 261  
Db 346 GAAGCAGATTTCAGAAAGTTCTCTCTCCACCATTTCTGGATCAGTTTACATCGTTACT 405  
QY 262 GGAATCTTGTAGGACAGATTCGTATCTGCTACTCCATGGACTTCTTTACTTACCGTT 321  
Db 406 GGACTATTAGTGTCTAACTCTTATATGTCCTCTGTGTTTACTATACTTACCAGTT 465  
QY 322 TCTACCTATTCCCTGATCTGTGCATCTCAGTTAGCCTTCAATGCTTTCTTTCTTTATTT 381  
Db 466 TCTACTTTCTCCTCATCTTGGCTCACAATTTGGCCTTCACTGCTTCTTCTCATATTTT 525  
QY 382 CTTAACCTCAGAAACCTTACCCCTATCATTTTAAATTCCTTTTCTCTTAACATAATCT 441  
Db 526 CTAACCTCGAGAAGTTACACCTTTCAATGTGAATTCCTGTTCTCTCTTACTATTTC 585  
QY 442 TCCACCTCTTGCATTCATATATAGGACAGACTCCACAAAAGTTACAAAAGGAG 501  
Db 586 CTGCCCCCTCTCTGGTGAACACTGATTCGGAACACACAGCAAGTGTCTAGAGTAA 645  
QY 502 TATGTCAAAAGTTTCATATGACCGTTGCTGCTGCTGCTGTTTATGCTTAGTCTTATCC 561  
Db 646 TATGTAGAGGATATATGATACCATTTGGTCTTCTGCTGGGATTTGGATTTGCTATCC 705  
QY 562 CTACACAGTAGCTTTCTAAAGTCTTAAGAGCAAAAATTTCTCAGAAATTATGGAT 621  
Db 706 CTGGTCAACTGATCTCTCAGGAAGTTTAAAGAGCAAAACATTTCTCAACGGTCACTGAC 765  
QY 622 ATGATAATCTAGTGTAGTGTAGGCCAGTTGTTAGCGTGGTGGGCTTTTTCGTAGC 681  
Db 766 TTGGTCGCTTACCAATCTCTAGTTGCAAGCTGTGGTTCTCATAGGACTTTTTCGCAAGC 825

QY 682 AGTGAATGGAACCTTTTGACGAGTGAATGGAATACTCAAAACATGGGAAGTATCTCTAC 741  
Db 826 GGGAGTGGAAACCTTTTAAACAGTGAATGGAATACTCAAACTGGGAAAGTGCCATAC 885  
QY 742 ATTATGAACCTAGTGTGGACAGCTGTACTCGCAGTATTCTCCATCGGTGCACAGGA 801  
Db 886 GTTATGACTTTGGCCTCGATAGCTATTCTTCGCAAGTCTACACCAATTGGCGTCTGTGG 945  
QY 802 CTGATCTTCGAGCTCTCTCTCTTATTCTCAAAATGCAATAAGCGTTTTCGGACTCC 861  
Db 946 CTGAATCTTGAAGTCACTTCTGTGTTCTCAATTCATTAATCTGCTGGGATTCCTATA 1005  
QY 862 GTTCTTATCTTGGCTGTAAATCAATTTCCATGACAAATGAATGGGTAAAGGTGATTTCT 921  
Db 1006 GTTCAGATTGTAGCAGTGATTGTTTCCATGATATAAATGAACGCGTCAAAAGATCTTCTCC 1065  
QY 922 ATGATCTTACCTATTGGGGTTTCACTTCTATGCTACCAACAATATCTTGATGACAAA 981  
Db 1066 ATCATTTTATGCTATCTGGGGATTTCATTTCTATCAGCACTACCTCGACGAAAAG 1125  
QY 982 AACTTGAAGAAAATCATGAAATCAACAACAGAAATCCCTGACCCACACAGACAGAA 1041  
Db 1126 AAGTTGAGACTAGCCACAAGTCTGTAGGAGATCTCTACTACTCTGCTGAGGAA 1185  
QY 1042 G 1042  
Db 1186 G 1186  
RESULT 10  
AX652015 3387 bp DNA linear PAT 22-MAR-2003  
LOCUS Sequence 908 from Patent WO03000898.  
DEFINITION AX652015  
ACCESSION AX652015  
VERSION AX652015.1 GI:29154833  
KEYWORDS Arabidopsis thaliana (thale cress)  
SOURCE Arabidopsis thaliana  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
1  
REFERENCE  
AUTHORS Chang, H.S., Chen, W., Cooper, B., Glazebrook, J., Goff, S.A., Hou, Y.M.,  
Katagiri, F., Quan, S., Tao, Y., Whitcham, S., Xie, Z., Zhu, T. and Zou, G.  
TITLE Plant genes involved in defense against pathogens  
JOURNAL Patent: WO 03000898-A 908 03-JAN-2003;  
Syngenta Participations AG (CH)  
FEATURES  
source Location/Qualifiers  
1..3387  
/organism="Arabidopsis thaliana"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:3702"  
ORIGIN  
Query Match 37.6%; Score 406.2; DB 6; Length 3387;  
Best Local Similarity 64.7%; Pred. No. 6.5e-89;  
Matches 622; Conservative 0; Mismatches 333; Indels 6; Gaps 1;  
QY 88 GGAGTATCTATGACAAACGGAGAAACAGTAATGCTTAGCAACGGTAGTTCACACTTGT 147  
Db 226 GTAGTTTACTATGAAATGTTGGGAGAGTACATGGATGGGAACACTTGTCCAACTAATC 285  
QY 148 GGCTTTCTGTGCTACTTCCATATTATCTTGTCTATTAAACACATGCAACACTGAT 207  
Db 286 GGCTTCCCTGTCTGTTTCTCTTCGCTTCTTTCCCAACCAAAATCCCAACCAACA 345  
QY 208 AGAGATGGAAAGAACCTCACCT-----AGAACCGGTGATGTGTTACGTAGTGGT 261  
Db 346 GAAGCAGATTTCAGAAAGTTCTCTCTCCACCATTTCTGGATCAGTTTACATCGTTACT 405  
QY 262 GGAATCTTGTAGGACAGATTCGTATCTGCTACTCCATGGACTTCTTTACTTACCGTT 321  
Db 406 GGACTATTAGTGTCTAACTCTTATATGTCCTCTGTGTTTACTATACTTACCAGTT 465  
QY 322 TCTACCTATTCCCTGATCTGTGCATCTCAGTTAGCCTTCAATGCTTTCTTTCTTTATTT 381  
Db 466 TCTACTTTCTCCTCATCTTGGCTCACAATTTGGCCTTCACTGCTTCTTCTCATATTTT 525  
QY 382 CTTAACCTCAGAAACCTTACCCCTATCATTTTAAATTCCTTTTCTCTTAACATAATCT 441  
Db 526 CTAACCTCGAGAAGTTACACCTTTCAATGTGAATTCCTGTTCTCTCTTACTATTTC 585  
QY 442 TCCACCTCTTGCATTCATATATAGGACAGACTCCACAAAAGTTACAAAAGGAG 501  
Db 586 CTGCCCCCTCTCTGGTGAACACTGATTCGGAACACACAGCAAGTGTCTAGAGTAA 645  
QY 502 TATGTCAAAAGTTTCATATGACCGTTGCTGCTGCTGCTGTTTATGCTTAGTCTTATCC 561  
Db 646 TATGTAGAGGATATATGATACCATTTGGTCTTCTGCTGGGATTTGGATTTGCTATCC 705  
QY 562 CTACACAGTAGCTTTCTAAAGTCTTAAGAGCAAAAATTTCTCAGAAATTATGGAT 621  
Db 706 CTGGTCAACTGATCTCTCAGGAAGTTTAAAGAGCAAAACATTTCTCAACGGTCACTGAC 765  
QY 622 ATGATAATCTAGTGTAGTGTAGGCCAGTTGTTAGCGTGGTGGGCTTTTTCGTAGC 681  
Db 766 TTGGTCGCTTACCAATCTCTAGTTGCAAGCTGTGGTTCTCATAGGACTTTTTCGCAAGC 825

Db 406 GCATATTAGTGTCTGCTAACTCTTATATATGTCTCTGTGGTTTACTATATCTTACCAGTT 465  
Qy 322 TCTACTATTCCCTGATCTGTGCATCTCAGTTAGCTTCAATGCTTTCTTCTCTTATTTT 381  
Db 466 TCTACTTTCTCCCTCATCTTGGCCCTCAAAATTTGGCCCTTCACTGCTTTCTTCTCATATTTT 525  
Qy 382 CTTAACTCACAAAACTTACCCTATCATTTTAAATTTCTTTTCTTCTTAACTATATCT 441  
Db 526 CTAACTCGCAGAAGTTACACCTTTCAATGTGAATTTCTGTCTTCTTACTATTTCC 585  
Qy 442 TCCACCTACTTGCAATTCATTAATGAGGAGACAGACTCCACAAAAGTTTACAAAAGGAGAG 501  
Db 586 TCTGCCCTCTCGTGTCAACTGATTCGGAACACAGCAAAAGTGTCTAGAGTAAA 645  
Qy 502 TATGTCAAGGTTTCATATGACCCCTTGTCTGCTGTCTGTTATGTTCTAGTCTTATCC 561  
Db 646 TATGTGATAGGATAATATGTAACCATTTGTTCTTCTGCTGGATTTGGATTCTGCTATCC 705  
Qy 562 CTACACAGCTAGCCCTTTTAAAGTCTTAAAGAGCAAAATTTCTCAGAAGTTATGGAT 621  
Db 706 CTGTGTACAACTGATCTCAGGAGGTTTAAAGAGCAAAATTTCTCAGCGTCACTGAC 765  
Qy 622 ATGATAATCTACGTGAGTCTAGTGCCAGTTGTGTAGCGTGGGCTTTTGTCTAGC 681  
Db 766 TTGGTCGCTTACCAATCTCTAGTTGCAAGCTGTGTGTTCTCATAGGACTTTTCGCAAGC 825  
Qy 682 AGTGAGTGGAAAACCTTTCGACGAGTGAATGATTAACACTCAACATGCGGAAGTATCCTAC 741  
Db 826 GGGGAGTGGAACCTTTTAAAGTGAGATGGAATACTCAAACTGGGGAAAAGTGCCTAC 885  
Qy 742 ATTATGAACCTAGTGTGGACAGCTCTTACCTGGCAGTTTATTTCTCATCGTGGCAGAGA 801  
Db 886 GTTATGACTTTGGCCCTCGATAGCTATTTCTGCGCAAGTCTACACATTTGGCGTGTGGA 945  
Qy 802 CTGATCTTCGAGCTCTCTCTCTATCTCAATGCAATGAAGCGTTTGGGACTCCCAAGT 861  
Db 946 CTGATCTTTGAGTCATCTCTGTGTCTTCCAAATTCATAAGCTGTGTGGATTGCTATA 1005  
Qy 862 GTTCTATCTTGGCTGATCATCTTTCCATGACAAATGAATGGTTAAAGTGATTTCT 921  
Db 1066 GTTCAGTGTGACGAGTGAATTTTCCATGATAAATAAGAACGCGTCAAGATCTTCTCC 1065  
Qy 922 ATGATCTTAGCTATTTGGGGTTTCACTTCTCTATGTCTACCAACATATCTTGTAGCAAAA 981  
Db 1066 ATCATTTAGCTATCTGGGATTCATTTCTATCTATCAGCACTACTCGACGAAAAG 1125  
Qy 982 AACTTGAAGAAAATCATGAATCAACACAGAAATCCCTGACCCACCAAGACAGAA 1041  
Db 1126 AAGTTGAAGACTAGCCACACAAAGTCTCTGAGGAGATCCTCATCTACTACTGCTGAGGA 1185  
Qy 1042 G 1042  
Db 1186 G 1186

RESULT 11  
AP006383/c  
LOCUS AP006383 73179 bp DNA linear PLN 22-JUL-2003  
DEFINITION Lotus corniculatus var. japonicus genomic DNA, chromosome 6,  
clone:LjT05120, TM0245, complete sequence.  
ACCESSION AP006383  
VERSION AP006383.1 GI:31581014  
KEYWORDS HTG  
SOURCE Lotus corniculatus var. japonicus (Lotus japonicus)  
ORGANISM Lotus corniculatus var. japonicus  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Lotaeae;  
Lotus.  
1  
REFERENCE  
Asamizu,E., Kato,T., Sato,S., Nakamura,Y., Kaneko,T. and Tabata,S.  
STRUCTURAL ANALYSIS OF A LOTUS JAPONICUS GENOME. IV. SEQUENCE  
FEATURES AND MAPPING OF SEVENTY-THREE TAC CLONES WHICH COVER THE

7.5 Mb Regions of the Genome  
DNA Res. (2003) In press  
2 (bases 1 to 73179)  
Sato,S.  
Direct Submission  
Submitted (07-MAY-2003) Shusei Sato, Kazusa DNA Research Institute,  
Department of Plant Gene Research; 2-6-7 Kazusa-Kamatari, Kisarazu,  
Chiba 292-0818, Japan (E-mail:ssato@kazusa.or.jp,  
URL:htp://www.kazusa.or.jp/, Tel:81-438-52-3935(ex.2337),  
Fax:81-438-52-3934)  
FEATURES  
Location/Qualifiers  
1. 73179  
/organism="Lotus corniculatus var. japonicus"  
/mol\_type="genomic DNA"  
/variety="japonicus"  
/db\_xref="taxon:34305"  
/chromosome="6"  
/clone="LjT05120"  
/clone\_lib="LjT library"  
/note="TAC clone:TM0245-synonym: Lotus japonicus"

## ORIGIN

Query Match 35.6%; Score 385.2; DB 8; Length 73179;  
Best Local Similarity 63.7%; Pred. No. 7.7e-84;  
Matches 604; Conservative 0; Mismatches 338; Indels 6; Gaps 1;  
Qy 68 CACATACAAACGGTGGCTCAGGAGTATCTATGCAACGGAGGAAACAGTAAATGGCTAG 127  
Db 64772 CAGCTACAAACCCCTTTTGGGAAGATTGTACTATGAAATGGTGAAGAGCAATGGATGG 64713  
Qy 128 CAACGGTAGTTCAACTTGTGGCTTTCCTGTGCTACTTCCATATATATATCTTGTCAATTA 187  
Db 64712 CAACACTGTGCCAGTGTCTGCTTCCAAATCTGCTGCCCTATATATACATCTCAGCAC 64653  
Qy 188 AAACACATGCAACAACTGATAGAGATGGAAGAAAGAACCTCACCTPAGAAACGGTGT- 242  
Db 64652 CCAAAGATCTCACTGAAAACAGCTTTTATCATCAGAATCAGCCTTCTAAATTAGTGCTAA 64593  
Qy 243 -ATTGGTTAGGTAGTGTCTGGACTTCTTGTAGGAGCAGATGTCTATCTGTACTCCATTG 301  
Db 64592 CACTTGTCTATGTCTCCCTTGGCCAACTTGTGACATTTAAATTTGTTATTTAAATTTCTGCTG 64533  
Qy 302 GACTTCTTTACTTACCCTGTTTCTACCTATTCCCTGATCTGTGCATCTCAGTTAGCCTTCA 361  
Db 64532 GGCTCTGGTCTCTCCAGTCTCTACTTATTCATCTATTTGTTTCACTCAGTTGGCTTCA 64473  
Qy 362 ATGCTTTTCTCTCTTATTTTCTTAACTCAAAAACCTTACCCCTATCATTTTAAATTTCTC 421  
Db 64472 ATGCTATTTTCTCTACTTCTCTCAATTCACCTCAAGATCACACCTTACATCAACTCTC 64413  
Qy 422 TTTTCCCTTAACTATATCTTCCACCTACTTGCATTTCAATATGAGGAGACAGACTCCA 481  
Db 64412 TAGTCCTTTAAACCAATTTCTCTGCTCTGCTTGTGTTTCAAGATGAGTCACCTGATTTCAA 64353  
Qy 482 CAAAAGTTTACAAAAGGAGAGTGTCTCAAAGTTTTCATATGCACCGTTGCTGCGTCTGCTG 541  
Db 64352 CACAAATCTCCACAAAAGTATGTCTATTGGATTCATTTACAGTTGCTGTCATCAGCTG 64293  
Qy 542 GTTATGCTGTAGTCTTATCCCTACAAACAGCTAGCCTTTCTAAAGTCTCTAAAGAGCAAA 601  
Db 64292 GGGTGGATTAGGACTTTCCCTTTACACAGTTTGTCTTCAAGAACGTCATAAAAAGGAAA 64233  
Qy 602 ATTTCTCAGAAGTTATGGATATGATATCTACGTGACTGTAGTGGCCAGTTGTGTTAGCG 661  
Db 64232 ATTTCAAAGTGTACATAGATATGATGTTTATCTGTCTCCCTTGTAGCTACATGTGCTAGTC 64173  
Qy 662 TGTGGGGCTTTTGTCTAGCAGTGTAGTGAAGAACTTTGAGCAGTGAATGGATAACTACA 721  
Db 64172 TAGTGGGACTTTTGTCTAGTGGAGAGTGAAGAGTTTGAAGAAAGAAATTTGAGGGGTATG 64113  
Qy 722 AACATGGGAAGGTATCCTACATTTATGAACCTAGTGTGGACAGCTGTTTACCTGGCAGTTAT 781  
Db 64112 AGACAGGTAAGCTGCCTATATATTTGAATCTTGTCTTTTACAGCCATAATTTTCGCAACTCG 64053

```
QY 782 TCTCCATCGGTGGCAGGAGTGAATCTTCGAGCTCTCCTCTCTATTTCTCAATGCAATAA 841
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
64052 TTAATCTCGGTGTGACGGGACTGGTTTCGAGGCTCTTCCCTCTTCTCCAATGCTATAA 63993
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 842 GCGTTTTGGGACNCCAGTGGTTCCTATCTTGGCTGTGAATCAATTTCCATGACAAAATGA 901
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
63992 GCGTTTTGGGTGACGATATGTTCCCTATATTAAGCTGTAAATGGTCTTACAGACAGATGC 63933
QY 902 ATGGGTTAAAGTGATTTCTATGATCCTAGCTATTTGGGGTTTCACTTCTCTATGTCTACC 961
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
63932 ATGGGATAAAGTCATTTCTATGCTGTGGCCGTTTGGGGTTTATATGCTAGCTGTATC 63873
QY 962 ACAAATATCTTGATGACAAAACCTTGAAGAAAATCATGAATACAA 1009
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
63872 AGCAATACCTGGATGAATGCAATCCAAACAGAGATAGAACAGAA 63825

RESULT 12
AX033548
LOCUS AX033548 1194 bp DNA linear PAT 21-SEP-2000
DEFINITION Sequence 5 from Patent WO0049152.
ACCESSION AX033548
VERSION AX033548.1 GI:10280292
KEYWORDS
SOURCE
ORGANISM
Arabidopsis thaliana (thale cress)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1
REFERENCE
AUTHORS Andre, B., Buerkle, L., Frommer, W. B. and Gillissen, B.
TITLE Nucleic acids that code for a nucleobase transporter
JOURNAL Patent: WO 0049152-A 5 24-AUG-2000;
ANDRE BRUNO (BE) ; BUERKLE LUKAS (DE) ; FROMMER WOLFGANG (DE) ;
GILLISSEN BERND (DE)
FEATURES
source
1. .1194
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"

ORIGIN
Query Match 34.2%; Score 369.6; DB 6; Length 1194;
Best Local Similarity 62.2%; Pred. No. 6.6e-80;
Matches 601; Conservative 0; Mismatches 359; Indels 6; Gaps 1;
QY 89 GAGTATCTATGACACGAGGAGGAAACAGTAATGGCTAGCAGCGTAGTTCAACTTTGTTG 148
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
197 GATTGTACTATGAAATGGAGGAAATAGCACATATGTGGTAACACTTCTTCAACTCATTTG 256
QY 149 GCTTTCTCTGTCTACTTCCATATATATCTGTCAATTTA-----AACACATGCAACAA 202
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
257 GCTTCCCTGTACTGTTCTGTTCCGGCTTCTTTCTCGAATCAGGCAACCAATCAACAG 316
QY 203 CTGATAGATGAAAGAAACCTCACCTAGGAACCGTGTATTTGGTTTACGTAGTCTTG 262
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
317 ATACAAATTTCACTAGTCAGTCCCTTCTCCACCACCTTCGATCGGTTTACTTGTGCACTG 376
QY 263 GACTTCTGTAGAGCAGATTTGCTGTACTCTCAATCGGACTTCTTTACTTACCGGTTT 322
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
377 GACTGTAGTGTCCGCTTATGCTTATTTGCTGCACTAGGGTTGCTTACTTACCAGTCT 436
QY 323 CTACCTATTCCCTGATCTGTCATCTCAGTTAGCTTCAATGCTTCTCTCTCTATTTTC 382
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
437 CTACTTCTCCCTCATCTTGGCCCTCAGCTGCGCTTCTACTGCTTTTCTCATATTTCC 496
QY 383 TTAACCTCACAACAACTTACCCCTATCATTTTAAATCTCTTTTCTCTTAACTATATCTT 442
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
497 TTAACCTCGCAAAAGTTCACTCTCTTGTAGTACTGTTCTTGTCTTCTCTCACTGTATCT 556
QY 443 CCACCCCTACTTGCATTTCAATATAGGAGACAGACTCCACAAAGATTACAAAAGGAGAT 502
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

```
Db 557 CTGCTCTCTTGTGTCACACTGATTCAGAAAACCTCAATATGATATCTAGAGTACAGT 616
QY | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
503 ATGTCAAAGGTTTCATATGACACCGTTCGTCGCTCTGCTGTTATGCTTAGTCTTATCCC 562
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
617 ATGTGATCGGTTTCATATGATACCATCGGTGCTTCGCTGGGATTTGACTGTTACTATCTC 676
QY | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
563 TACAACAGCTAGCCTTTCTAAAGTCTTAAAGAGCAAAATTTCTCAGAGTTATGGATA 622
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
677 TGATACAAATGCTCTTTCAGGAAAGTTTTCACGAAGCATACATCTCAGCAGTCACGACT 736
QY | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
623 TGATATCTACGTGAGTCTAGTGGCCAGTTGTTAGCGTGGTGGGCTTTTGTCTAGCA 682
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
737 TGCCATTTTACCACTCTCTAGTTGGAGTTGTAGTTCTCATAGGACTTTTTTGCAGGTG 796
QY | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
683 GTGAGTGGAAAACCTTTGAGCAGTGAATGGAATACTACAAACATCGGAAAGTATCTTACA 742
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
797 GAGAGTGGGAACTTTGCCAGTGAGATGAGAACTTACAACTCGGAAAGTGTCTATATG 856
QY | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
743 TTATGAACCTAGTGTGGACAGCTTTTACCTGCGAGTTATTTCTCCATCGGTGGCAGGAC 802
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
857 TTTTGACTTTTAGCTCGGAGCTATTTCTCTGCAAGTCTACACTCTTGTCTTGTGGAT 916
QY | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
803 TGATCTTTCGAGCTCTCTCTCTATTTCTCAATGCAATAAGCGTTTGGGACTCCCGAGTGG 862
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
917 TGATCTTCGAGTCATCTCTGTTCTTCCAATTCATTAACAGCTGTGGGATTTGCCTATAG 976
QY | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
863 TTCTCTATCTTGGCTGTAATCAATTTTCCATGACAAATGAATGGTTTAAAGGTTATTTCTA 922
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
977 TTCCAGTTGCGGAGCTAGTTTTCATGATGAGATGAGCGCATCAAAAATCTTCTTCCA 1036
QY | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
923 TGATCTAGCTATTTGGGTTTCACTTCTATGCTTACCAACAATATCTTGTATGACAAA 982
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
1037 TTAATTTAGTATCTCGGCTTCTTCTTCTATCTGCTCTATCAGCACTACCTCGACGAAAAGA 1096
QY | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
983 ACTTGAAGAAAATCATGAATCAACAACAGATCCCTGACCCACCAAGAGAGAG 1042
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
1097 AGTTGAATACTAGCCACACAAAGTGTCTAGGAGATCTTCTATCTACCTGTTAGGAAGGTC 1156
QY | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
1043 AGTCAA 1048
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
1157 ACACAA 1162

RESULT 13
AX096558
LOCUS AX096558 1165 bp mRNA linear PLN 18-SEP-2002
DEFINITION Arabidopsis thaliana putative protein (At4g18200) mRNA, complete cds.
ACCESSION AY096558
VERSION AY096558.1 GI:20465496
KEYWORDS FLI CDNA.
SOURCE
ORGANISM
Arabidopsis thaliana (thale cress)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 1165)
Yanada, K., Ban, J., Chan, M. M., Chang, C. H., Chang, E., Dale, J. M.,
Deng, J. M., Goldsmith, A. D., Lee, J. M., Onodera, C. S., Quach, H. L.,
Tang, C., Toriumi, M., Wu, H. C., Yamamura, Y., Yu, G., Bowser, L.,
Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J.,
Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B.,
Lin, J., Meyers, M. C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C. J.,
Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A.,
Shinozaki, K., Davis, R. W., Ecker, J. R. and Theologis, A.
Arabidopsis Open Reading Frame (ORF) Clones
2 (bases 1 to 1165)
Yanada, K., Ban, J., Chan, M. M., Chang, C. H., Chang, E., Dale, J. M.,
Deng, J. M., Goldsmith, A. D., Lee, J. M., Onodera, C. S., Quach, H. L.,
Tang, C. C., Toriumi, M., Wu, H. C., Yamamura, Y., Yu, G., Bowser, L.,
Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J.,
Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B.,
```

Lin, J., Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A., Shinohara, K., Davis, R.W., Ecker, J.R. and Theologis, A.

#### TITLE JOURNAL

Submitted (16-APR-2002) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA  
The RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL CDNA : 'RIKEN Arabidopsis Full-length cDNA'): Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinohara, K.

#### COMMENT

The Salk, Stanford, PGE (SSP) Consortium members constructed and sequenced the PUNI (ORF) clones using the RAFL cDNAs: Yamada, K., Bath, J., Chan, M.M., Chang, C.H., Chang, E., Dale, J.M., Deng, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Bowser, L., Chen, H., Cheuk, R., Jones, T., Karlin-Neumann, G., Kim, C., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Nguyen, M., Palm, C.J., Shinn, P., Southwick, A., Davis, R.W., Ecker, J.R. and Theologis, A.

Yamada, K. (SSP/PGE) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinohara, K. (RIKEN GSC) and Theologis, A. (SSP/PGE) contributed equally to this work as PIs.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

#### FEATURES source

Location/Qualifiers  
1..1165  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
/chromosome="4"  
/clone="U18628"  
/ecotype="Columbia"  
/note="This clone is in pUNI 51."  
1..1165  
/gene="At4g18200"  
1..1134  
/gene="At4g18200"  
/codon\_start=1  
/evidence="experimental"  
/product="putative protein"  
/protein\_id="At4g18200.1"  
/db\_xref="GI:20465497"  
/translation="MGISQVHYCNGDQNLNLANLLDHETESFVPTQKNCRWLRVSI  
YAFVFCOPLATVGRYVYGGKSTVYVLLQIGPVLILFRFSRIRQPKSTDT  
NFSQSPFTLLASVLCGLVAVLSAVGLLYLPVSTFSLILASQIAPAFSVF  
LNSQKFTPLVNSLFLVSSALLVNTSDNTNVRQVQVIGFICTIGASAGIGLV  
LSLQLLFRKVFTRHTSSAVIDLANYQSLVATCVVLGLFASGWRILPSEMRNYKLG  
KVSYLTLASAAIFWQVYTVGCVGLIFESSVFNSTITAVGLPIPVVAVIVFHKMD  
ASKFSIILAIWGLSFVYQHYLDEKLLKCTQKRVBEETQL"  
1135..1165  
/gene="At4g18200"

#### 3'UTR

#### ORIGIN

Query Match 34.1%; Score 368.2; DB 8; Length 1165;  
Best Local Similarity 62.8%; Pred. No. 1.5e-79;  
Matches 590; Conservative 0; Mismatches 343; Indels 6; Gaps 1;  
QY 59 TCACCTACACATCAACGGTGCCTCAGGAGTACTATGACACGAGGAAACAGTA 118  
DB 152 CTGCGAACCACTTCTCAGTTCCTGGGTAGACTACTATGAAATGGAGGAAACGA 211  
QY 119 AATGCTAGCAACGGTAGTTCAACTTGTGGCTTTCCTGTGCTACTTCCATATTATCT 178  
DB 212 CATATGTGTTACACTTCTTCACTCACTTGGCTTCCCTGTACTGATTCTCTCGCTTCT 271  
QY 179 TGTCTATTTA-----AACAACATGCAACACTGATAGAGATGGAAAAAGACCTCACCTA 232  
DB 272 TTTCTCGAATCAGGCAACCAATCAACAGATACAAATTTTCACTCAGTCCCTCTCTTCA 331  
QY 233 GGAACCGGTATTTGGTTTACGTAGTGTGGACTTCTTGTAGGACGAGATTGCTATCTGT 292

Db 332 CCACCTTGGCATCGTTTACTTGTGCACTGAGTGTCTGCTTATGCTTATTTGT 391  
QY 293 ACTCCATGGACTCTTTTACTTACCCGTTTCTACCTATTCCTGATCTGTGCACTCAGT 352  
DB 392 CTGCACTGGGTGCTTACTTACCACTCTACTTCTCCCTCATCTTGGCCCTCAGT 451  
QY 353 TAGCCTTCAATGCTTCTTCTTATTTTCTTAACTACACAAAACACTTACCCCTATCATTT 412  
DB 452 TGGCCTTCACTGCTTCTTCTATATTTCTTAACTCGCAAAAGTTTCACTCTTTGATAG 511  
QY 413 TAAATCTCTTTTCTTAACTATATCTTCCACCTTCTTAACTCAATAATGAGGAGA 472  
DB 512 TCAATCTTGTGTTCTTACGGTTTCTTCCCTCTGCTCTGCTGCTCAACACTGATTCAG 571  
QY 473 CAGACTCCACAAAGTTTACAAAAGGAGATGATGTAATCTACGTGAGTCTAGTGGCCAGTT 532  
DB 572 AAAAAACAACAAATGATCTAGAGTACAAATGATGATGATGATGATGATGATGATGATG 631  
QY 533 CGTCTGCTGTTATGTTCTAGTCTTATCCCTACACAGCTAGCTTCTTAAAGTCCCTAA 592  
DB 632 CTTCCGCTGGGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 691  
QY 593 AGAAGCAAAATTTCTCAGAAAGTTATGATGATGATGATGATGATGATGATGATGATGATG 652  
DB 692 CGAAGCATACATCTCAGCAGTCTTGGCACTTGGCACTTGGCACTTGGCACTTGGCACTT 751  
QY 653 GTGTTAGCGTGGGGCTTTTCTAGCAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAG 712  
DB 752 GTGTGGTACTCATAGGACTGTTTGCAGAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGA 811  
QY 713 ATAACTCAACATGGAAGTATCTTACATGATGATGATGATGATGATGATGATGATGATGATG 772  
DB 812 GAAACTCAAACTGGGAAAGTGTATATATCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTT 871  
QY 773 GGCAAGTATACATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 832  
DB 872 GGCAAGTATACATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 931  
QY 833 ATGCAATAAGCGTTTGGGACTCCAGTGGTCTTCTATCTTGGCTGTTGTTGTTGTTGTTGTTGTT 892  
DB 932 ATTCATAAATGCTGTGGGACTGCTTATAGTTCAGCTTGTAGCAGTGTAGTGTAGTGTAGTGT 991  
QY 893 ACAAATGAATGGTTAAAGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 952  
DB 992 ATAAGATGGATGCAATCAAGATTTTCTCCATCATTTTAGCTATCTGGGGCTTCTTCTTCT 1051  
QY 953 ATGCTTACCACAAATATCTTGTGATGACAAAAACTTTGAAGA 991  
DB 1052 TTGCTATCAGCACTACTCTCGACGAAAGAGTTGAAGA 1090

#### RESULT 14

#### LOCUS

AY074546 1428 bp mRNA linear PLN 18-SEP-2002  
Arabidopsis thaliana unknown protein (At4g18200) mRNA, complete cds.

#### ACCESSION

AY074546.1

#### VERSION

AY074546.1

#### KEYWORDS

FLI CDNA.

#### SOURCE

Arabidopsis thaliana (thale cress)

#### ORGANISM

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

1 (bases 1 to 1428)

#### REFERENCE

#### AUTHORS

Yamada, K., Liu, S.X., Sakano, H., Pham, P.K., Banh, J., Chung, M.K.,

Goldsmith, A.D., Lee, J.M., Quach, H.L., Toriumi, M., Yu, G., Bowser, L.,

Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J.,

Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B.,

Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T.,

Satou, M., Seki, M., Shinn, P., Southwick, A., Shinohara, K.,

Davis, R.W., Ecker, J.R. and Theologis, A.

TITLE JOURNAL REFERENCE AUTHORS	Arabidopsis Full Length cDNA Clones Unpublished 2 (bases 1 to 1428) Yamada, K., Banh, J., Chan, M.M., Chang, C.H., Chang, E., Dale, J.M., Deng, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Bowser, L., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A. Direct Submission Submitted (16-JAN-2002) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA
TITLE JOURNAL COMMENT	RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAPL cDNAs (RAPL cDNA: 'RIKEN Arabidopsis Full-length cDNA'): Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.  The Salk, Stanford, PGEC (SSP) Consortium members carried out the sequencing and annotation of the RAPL cDNAs: Yamada, K., Banh, J., Chan, M.M., Chang, C.H., Chang, E., Dale, J.M., Deng, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Bowser, L., Chen, H., Cheuk, R., Jones, T., Karlin-Neumann, G., Kim, C., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Nguyen, M., Palm, C.J., Shinn, P., Southwick, A., Davis, R.W., Ecker, J.R. and Theologis, A.  Yamada, K. (SSP/PGEC) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP/PGEC) contributed equally to this work as PIs.  Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.
FEATURES source	Location/Qualifiers 1..1428 /organism="Arabidopsis thaliana" /mol_type="mRNA" /db_xref="taxon:3702" /chromosome="4" /clone="RAF109-36-N15 (R18628)" /ecotype="Columbia" /note="This clone is in a modified pBluescript vector (FLC-1) as a BamHI/XhoI insert." 1..1428 /gene="At4g18200" 1..71 /gene="At4g18200" 72..1205 /gene="At4g18200" /evidence=experimental /product="unknown protein" /protein_id="AA169512.1" /db_xref="GI:18491221" /translation="MGISOVHYCNGDONLEANLLDHEETESVPQTKCKEKLRSVI YAFVFCOPLATVGLRXYENGKSTVYVLLQIGPVLJLFRFSRIRPKSTDT NFSQSPSFTLASVLCGLLSAVAYLSAVGLLYLPVSTFSLIASQLAFTAFSYF LNSQKPLIVNSLFLTVSSALLVVDNTSENTNVSRVQYVIGFICTIGASAGLV LSLIQLFRKVTFTKSSAVLDLANYQSLVATCVLLIGLFASGEWRITPSEMRNYKLK KVSYLITLASAIFWQVTVGCVGLKIFSSSVFSNSITAVGLPIPVVAVIVFDKMD ASKIFSIIILAIWGFLSFVYQHYLDEKLLKCTQKPEBETQTL" 1206..1428 /gene="At4g18200" misc_difference 1412 /gene="At4g18200" /note="not present in genomic sequence"
ORIGIN	Query Match 34.1%; Score 368.2; DB 8; Length 1428; Best Local Similarity 62.8%; Pred. NO. 1.4e-79; Matches 590; Conservative 0; Mismatches 343; Indels 6; Gaps 1;  RESULT 15 AK071254 LOCUS DEFINITION Oryza sativa (japonica cultivar-group) cDNA clone:J023087G23, full insert sequence. AK071254 ACCESSION VERSION AK071254.1 GI:32981277
3' UTR	1123 TTGCTATCAGCACTACCTCCAGCAAGAAAGAGTTGGAAGA 1161
misc_difference 1412	
gene	59 TCACCTCTACACATACAAACGGTGGCTCAGGAGTATATCTATGACAAACGAGGAAACAGTA 118 223 TCTGCCAACCACTTGTCTACAGTTCTGGGTAGACTGTACTATGAAATGAGGGAAGCA 282 119 AATGGCTAGCAACGGTAGTTCACATTTGGTGGCTTTCCTGTGCTACTTCCATATATATCT 178 283 CATATGTGGTAACACTTCTTCAACTCAATGGCTTCCCTGTGACTGATCTCTTCCGCTTCT 342 179 TGTCTATTTA-----AAACACATGCAACAACTGATGAGATGGAAGAAACCTCACCTA 232 343 TTTCTCGAATCAGGCAACCCAAATCAACAGATACAAATTTTCAGTCAGTCCCCCTTCCTTCA 402 233 GGAACCGTGTATTTGGTTTACGTTAGTGTCTTGGACTTCTTGTAGGAGCAGATGCTATCTGT 292 403 CCACCTTGCATCGGTTTACTTGTGCACCTGGACTGCTAGTGTCTGCTTATGCTTATTTCT 462 293 ACTCATTTGGACTTCTTTACTTACCGGTTTCTACCTATTCCTCTGATCTGTGATCTCATGT 352 463 CTGCAGTTGGGTTGGCTTTACTTTACCAGTCTCTACTTCTCCCTCATCTTGGGCTCACAGT 522 353 TAGCCCTTCAATGCTTTCTTCTCTTATTTCTTAACCTCAACAAACCTTACCCCTATCATTT 412 523 TGGCCTTCACTGCCCTTTTCTCATATTTCTTAACTCGCAAAAGTTTCACTCTCTTTGATAG 582 413 TAAATTTCTTTTCTCTTAACTATATCTTCCACCCTACTTGCATTTCAATAATAGGAGA 472 583 TCAATTTCTTTTCTCTTACGGTTTCTCTGCCCTCTCTGCTGCTCAACACTGATTTCAG 642 473 CAGACTCCACAAAGTTACAAAGGAGATATGTCAAAAGTTTTCATATGACACCGTTGCTG 532 643 AAAACACAAACAAATGTATCTAGAGTACAATATGTATGGGTTCATCTGTACCATTTGGT 702 533 CGTCTGCTGGTTATGGTCTAGTCTTATCCCTACACAGCTAGCTCTTCTTAAAGTCTTAA 592 703 CTTCCGCTGGGATTGGATTGGTGTACTATCTCTGATACAACTGCTCTTCAGGAAAGTTTCA 762 593 AGAAGCAAAATTTCTCAGAAGTTATGGATATGATAATCTACGTGAGTCTAGTGGCCAGTT 652 763 CGAAGCATACATCCTCAGCAGTCTCTGGACTTGGCCAAATTACAGTCTCTAGTTGCAACTT 822 653 GGTGTAGCGTGGTGGGCTTTTGTGTAGCAGTGTGGAACACTTTTGAGCAGTGAATAAG 712 823 GTGTGCTACTCATAGGACTGTGTGCAAGTGGAGAGTGGAGAACTCTGCCAAGTGAGATGA 882 713 ATAACCTACAAACATGGGAGGTATCCTACATTTATGAACCTAGTGTGGACAGCTGTACTCT 772 883 GAAACTACAAACTGGGGAAGGTTCATATATCTTGACTTTGGCCCTCAGCAGCTATTTTCT 942 773 GGCAGTTATTTCTCCATCGTGGCAGCAGGACTGATCTTCGAGCTCTCTCTCTTATTTCTCAA 832 943 GGCAGTATATACATGTTGTTGTGGGATTAATCTTCAGTCTTCTCTGTGTTCTCCA 1002 833 ATGCAATAGCGTTTGGGACTCCAGTGGTTCCTATCTTGCTGTAATCATTTTCCATG 892 1003 ATTCCATAACTGCTGTGGGACTGCCTATAGTTCAGTTTGTAGCAGTGTAGTTTCCATG 1062 893 ACAAAATGAATGGGTAAAGTGATTTCTATGATCCTAGTATTTTGGGGTTTCACTTCTCT 952 1063 ATAAGATGGATGCATCAAAAGATTTTCTCCATCATTTTATAGTATCTGGGGCTTCTCTTCT 1122 953 ATGCTTACCAACAATATCTTGTATGACAAAAACTTCAAGA 991 1123 TTGCTATCAGCACTACCTCCAGCAAGAAAGAGTTGGAAGA 1161



## KEYWORDS

Oryza sativa (japonica cultivar-group)

## SOURCE

Oryza sativa (japonica cultivar-group)

## ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.

## REFERENCE

## AUTHORS

1  
The Rice Full-Length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-Length cDNA Project Team; Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuka, K., Shishiki, T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group; Otomo, Y., Murakami, K., Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y., Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M., Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J., Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J., Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN; Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Osato, N., Ota, Y., Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Yoshino, M., and Hayashizaki, Y.

## TITLE

Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice

## JOURNAL

Science 301 (5631), 376-379 (2003)

## MEDLINE

22752273

## PUBMED

12869764

## REFERENCE

## AUTHORS

2 (bases 1 to 1476)  
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K., Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imamura, K., Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M., Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M., Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A., Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M., Namiki, T., Narikawa, R., Niikura, J., Nishi, K., Nomura, K., Numasaki, R., Ohneda, E., Ohno, M., Ohtsuka, K., Oka, M., Ooka, H., Osato, N., Ota, Y., Otomo, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K., Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akashira, S., Tanaka, T., Tomaru, A., Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W., Yanada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S., and Yoshimura, A.

## TITLE

Direct Submission

## JOURNAL

Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan [E-mail: skikuchi@ias.affrc.go.jp]

Tel:81-29-838-7007, Fax:81-29-838-7007

## COMMENT

This clone is one of the 28K full-length cDNA clones from japonica rice.

URL : <http://cdna01.dna.affrc.go.jp/cDNA/>

NIAS Rice Full-Length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuka, K., Shishiki, T., and Yamamoto, M.

## REFERENCE

FAIS Genome Sequencing & Analysis Group; Otomo, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Narikawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K., and Murakami, K.  
Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K.,

Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J., Kouda, M., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Osato, N., Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akashira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K., Yasunishi, A., and Hayashizaki, Y.

FEATURES  
Location/Qualifiers  
source  
1..1476  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="mRNA"  
/culturvar="Nipponbare"  
/db\_xref="taxon:39947"  
/clone="J023087G23"

## ORIGIN

Query Match 28.4%; Score 307; DB 8; Length 1476;  
Best Local Similarity 59.4%; Pred. No. 1.5e-64;  
Matches 545; Conservative 0; Mismatches 360; Indels 12; Gaps 1;

Qy 92 TATATATGCAACGGAGGAAACAGTAAATGGCTAGCAACGGTAGTTCAACTTGTGGCT 151  
Db |||||  
Qy 326 TCTACTACAATCAAGCGGCAATAGCAAGTGGATGCCAATTCGTCCAAACTGCTGCT 385  
Db |||||  
Qy 152 TTCC-----TGTGCTACTTCCATATATATATCTTTGTCTATTTAAACACATGCAA 199  
Db |||||  
Qy 386 TTCCGATTTTGTTCATTTGCCCTATTCTTTTCCATTCAAAGACATCTTCTACACAAACAG 445  
Qy 200 CAATGATAGAGATGGAAGAAGAACCTCACCTAGGAACCGGTATTTGGTTAGCTAGTGC 259  
Db |||||  
Qy 446 TCACCTAGTAGTCTCGCCCTACAAATTTCTATCCCAAAATTAATCTCTGATATATGTTGCC 505  
Qy 260 TTGGACTCTTGTAGGAGCAGATTTGCTATCTGTACTTCCATTGGACTTCTTTTACTTACCCG 319  
Db |||||  
Qy 506 TGGGCCCTCATCATTTGCTGAGACGACTTGGTGTATTTCTTATGGCTACTATATCTTCGG 565  
Qy 320 TTTCTACTATTCCTCGATCTGTGCATCTCAGTTAGCCCTTCAATGCTTTCTTCTCTTATT 379  
Db |||||  
Qy 566 TCTCAACATATTGCTCATCTGTGCTAGTCAGCTTGGCTTCAATGCTGTCTTCTCATATT 625  
Qy 380 TTCTTAACACAAAAAATTTACCCCTATCATTTTAAATTTCTTTTCTTCTTAACTATAT 439  
Db |||||  
Qy 626 TCCTCAATGCTCAAAAAATTCACCCCTCTGATTTTCAATTCGGTAGTCTCTCTTACGTTTT 685  
Qy 440 CTTCCACCCTACTTTGCATTCAATAATGAGGAGCAGACTCCACAAAAGTTACAAAAGGAG 499  
Db |||||  
Qy 686 CTGCTTCACTCTTGGAGTTGATGAAGATTCTCAGGGAACCTACTAGTATATCATCGGA 745  
Qy 500 AGTATGTCAAAGGTTTCATATGCAACCGTTGCTGCGTCTGCTGGTTATGGTCTAGCTTAT 559  
Db |||||  
Qy 746 AGTACATTTTGGGTTTCTCTGTGACACTAGGGGCATCAGCTACATCTCGCTCATTTCTCT 805  
Qy 560 CCCTAACACAGCTAGCCTTTCTTAAAGTCTTAAAGAGCAAAATTTCTCAGAAAGTTATGG 619  
Db |||||  
Qy 806 CCCTGATGCAAGTCACATTTGAGAAGGTTATTAAAGAGGAGACCTTCTCAGTTGTGTGA 865  
Qy 620 ATATGATAAATCTACGTGAGTCTAGTGGCCAGTTTGTGTAGCGTGGTGGGCTTTTGGCTA 679  
Db |||||  
Qy 866 ACATGCGATATATACAGCTCTCTGTGCAACATTTGGCTTCTTCTTGGTGGGTTATTTGCA 925  
Qy 680 GCAGTGAAGTGAAGAACTTTGAGCAGTAGTGAATGGATAACTACAAACATGGGAAGTATCCT 739  
Db |||||  
Qy 926 GTGGTGAATGGATGACTTTTACAAGAGAGAGATGCAATCCAGTCTGGGAAGCTGTCAT 985  
Qy 740 ACATTATGAACCTAGTGTGGACAGCTGTTACCTGGCAGTTATTTCTCCATCGGTGGCAG 799  
Db |||||  
Qy 986 ATGTAATGACACTGCTGTGGACGGCTATATCTTGGCAGGTAGCATCAGTTGGAGTGGTGG 1045



```
Qy      800 GACTGATCTTCGAGCTCTCCTCTATCTCAAAATGCAATAGCGTTTGGGACTCCCAG 859
Db      1046 GATTGATCTTTGTGGTGTTCATCGCTGTTTCAAATGTGATAGCACCTAGCTCTACCCA 1105
Qy      860 TGGTTCTATCTTGGGCTGTAATCATTTCCATGACAAATGAATGGTTAAAGGTGATTT 919
Db      1106 TCATTCTGTTTTGCTGTGATTTCTTCATGACAAGATGGATGGAGTAAAGATTATAG 1165
Qy      920 CTATGATCCTAGCTATTTGGGGTTTCACCTTCTCTATGCTACCAACAATATCTTTGATGACA 979
Db      1166 CTATGCTGATGGCCATTTGGGGATTTATGTCATATGCCCAATATATGTTGATGGCA 1225
Qy      980 AAACTTGAAGAAAAAT 996
Db      1226 AGAAGGTAGAAAGACT 1242
```

Search completed: November 1, 2004, 15:44:48  
Job time : 4796.3 secs

**THIS PAGE BLANK (USP 107)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 1, 2004, 02:40:28 ; Search time 560.296 Seconds  
(without alignments)  
9996.736 Million cell updates/sec

Title: US-09-913-767-10

Perfect score: 1067  
Sequence: 1 agacaagaatggtgaaggct.....gaccataagtgaagttaa 1067

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_23sep04:\*

1: Geneseqn1980s:\*

2: Geneseqn1990s:\*

3: Geneseqn2000s:\*

4: Geneseqn2001as:\*

5: Geneseqn2001bs:\*

6: Geneseqn2002as:\*

7: Geneseqn2002bs:\*

8: Geneseqn2003as:\*

9: Geneseqn2003bs:\*

10: Geneseqn2003cs:\*

11: Geneseqn2003ds:\*

12: Geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	476.4	44.6	1049	3	AAA97920 A. thalia
2	428.4	40.1	1225	3	AAA97919 A. thalia
3	342.8	32.1	2586	3	AAC44184 Arabidops
4	177.4	16.6	493	3	AAC36831 Arabidops
5	147	13.8	1071	3	AAA97925 A. thalia
6	145.6	13.6	3387	6	ABZ14407 Arabidops
7	145.6	13.6	3387	8	ADA68246 Arabidops
8	141.6	13.3	1152	8	ADA70213 Rice gene
9	136.2	12.8	1194	3	AAA97923 A. thalia
10	135	12.7	1047	6	ADG88190 A. thalia
11	135	12.7	1081	3	AAA97924 A. thalia
12	131.4	12.3	1200	8	ADA70627 Rice gene
13	119.8	11.2	1293	3	AAA97922 A. thalia
14	116.6	10.9	1145	3	AAA97921 A. thalia
15	116.6	10.9	1292	3	AAC51119 Arabidops
16	116.6	10.9	1295	3	AAC37289 Arabidops
17	114	10.7	83698	6	ABN85767 Arabidops
18	112.8	10.6	2660	3	AAC47974 Arabidops
19	112.6	10.6	1403	3	AAC37652 Arabidops
20	107.6	10.1	1418	3	AAC46395 Arabidops
21	107.6	10.1	1421	3	AAC39779 Arabidops

22	81.6	7.6	2175	8	ADA70748	Ada70748 Rice gene
23	72	6.7	592	10	ABX56692	Abx56692 Arabidops
24	51.6	4.8	1155	8	ADA69686	Ada69686 Rice gene
25	49.8	4.7	2000	8	ADA71938	Ada71938 Rice gene
26	47	4.4	774	6	ABN99127	ABn99127 Arabidops
27	46.8	4.4	477	3	AAC37273	Aac37273 Arabidops
28	42.2	4.0	41104	6	AAD36260	Aad36260 Human G-p
29	41	3.8	1040	8	ADA69568	Ada69568 Rice gene
30	40.6	3.8	6317	6	ABL32408	Ab132408 Human inn
31	40.6	3.8	6317	6	ABL49311	Ab149311 Human pol
32	39.4	3.7	9974	2	AAX20573	Aax20573 Polynucle
33	39.2	3.7	478	9	ACH35469	Ach35469 Human end
34	39.2	3.7	896	8	ABX05058	Abx05058 Human nov
35	39	3.7	110000	6	ABN71527_06	Continuation (7 of
36	38	3.6	2000	8	ADA71938	Ada71938 Rice gene
37	38	3.6	6621	6	ABK33933	Abk33933 Human DNA
38	38	3.6	6621	6	ABL70156	Ab170156 Chemical
39	38	3.6	6621	6	AAS61098	AA61098 Human gen
40	38	3.6	6621	8	ABZ09984	Abz09984 Haematopo
41	38	3.6	6621	8	ABZ10098	Abz10098 Haematopo
42	38	3.6	6621	8	ADA20345	Ada20345 Prostate
43	38	3.6	6621	8	ADA84152	Ada84152 Human ren
44	38	3.6	6621	10	ADE84078	Ade84078 Human lym
45	37.6	3.5	1024	6	ABX65613	Abx65613 Helicobac

ALIGNMENTS

RESULT 1  
AAA97920  
ID AAA97920 standard; DNA; 1049 BP.  
AC AAA97920;  
XX  
XX  
DT 19-JAN-2001 (first entry)  
XX  
DE A. thaliana PUP1 DNA #2.  
XX  
XX  
KW PUP1; transgenic plant; nucleobase transporter; apical dominance;  
KW flowering behaviour; senescence; pesticide distribution; ds.  
XX  
OS Arabidopsis thaliana.  
XX  
XX  
PN DE19907209-A1.  
XX  
PD 24-AUG-2000.  
XX  
XX  
PF 19-FEB-1999; 99DE-01007209.  
XX  
XX  
PR 19-FEB-1999; 99DE-01007209.  
XX  
XX  
(FROM/) FROMMER W.  
XX  
XX  
Gillissen B, Buerkle L, Andre B, Frommer WB;  
XX  
XX  
WPI; 2000-566202/53.  
XX  
XX  
PT Nucleic acid, useful for producing transgenic plants with altered  
PT nucleobase transporter, encodes a nucleobase transporter protein of  
PT Arabidopsis thaliana.  
XX  
XX  
PS Claim 2; Page 12; 24pp; German.  
XX  
XX  
This invention describes a novel nucleic acid encoding a plant nucleobase  
transporter (I). (I) is produced by complementation of a nucleobase  
transport (NBT)-defective host cell with a plant gene bank by selection  
of NBT-positive cells. (I) is used to isolate homologous sequences from  
bacteria, fungi, plants, animals and humans, for expression of the  
encoded protein (II) in prokaryotic or eukaryotic cells, for inhibiting  
expression of (II) (when in antisense orientation), and to produce  
transgenic crop plants. The transgenic plants have modified nucleobase  
transport properties, e.g. altered affinity and substrate specificity



Db 77 GTGAGGTCCTTTGTTAACTCGTCTCTACTTCAACCAATGCGGAAACGAATCTGTTCA 136  
Qy 121 CTACGTTTCTTGAACCTGAGGCTTTCCTGTATCTTCACTCTCTGCTCTCTCTTACA 180  
Db 137 TGAGCTTCTATCAACCGCTGGTTTCCATCATCTCTCTCTCTCTCTCTCTCTTCC 196  
Qy 181 TTACCCGCGCGAGAACAAATGTGGTGA-----TAGTACAAGTT 222  
Db 197 TCAGCGTCCGCGCAACCGCAACCTTAACAACGCGGAAACAAAGCGGAAACAAAGC 256  
Qy 223 TCTTCTTATCAACCGCTCTCTTATCGCGCTGTTATGTAGGCATCTCTCAGGT 282  
Db 257 TCTTCTCTATGGAACCTCTCTGTTATCGCTCCATTTGTATAGGGTTGCTCACAGGAC 316  
Qy 283 TTGATAACTACTTGATCATATGGTATAGCTTATCTTCAGTTTCTACAGCTGCTCTTA 342  
Db 317 TTGACAACTACTTATATCTTACGGATTAGCATATCTGCAAGTTTCACTTATCGCTCA 376  
Qy 343 TCATTGCTTTCAGAGTGTATTTATAGTATCTTCTCATTTCTTATGTTTAAACATAAAGT 402  
Db 377 TAATCGAACTCAACTAGCTTTCAACGCTCTCTCTGCTTTCTTGTAGTCAAGCAAACT 436  
Qy 403 TCACCTCTTTTACCATCAATGCTGTTGTTGTGTTGATCTGTTGCTGCGGTTTGGGAA 462  
Db 437 TCACCTCGGTTCTCCATAAACGCGCTGCTTTTGTGACGGTTGCTATCGGGATCCTTGC 496  
Qy 463 TGCATACCGAACTGATAGCCAGTTTCATGAGACTCACAAGCAGTACATAAAGTCTTCT 522  
Db 497 TACACAGTATGAGGACAAACCGCTAAGGAGGAGCAAGAAAGATATGTTGGGTTCT 556  
Qy 523 TGATTACTGTAGCAGCTGTTATGATGCTTTTCACTTGGCAATTTAGTGAAGTCTGTT 582  
Db 557 TGATGACTGTGGTTGAGCTCTTCTCTATGCTTTTATATTTACGCTCGTTGAGTAACTT 616  
Qy 583 ACCAGAAAGCTAAGCAACCATGAGCTATACCTTGTGCTCGAGTTCAGTTGATTTGTT 642  
Db 617 ACAAGAAAGCTCGTCAAGAAATCACTTCCCACTTGTGCTTGGATTCAGATGGTCA 676  
Qy 643 GTCTCTCTCTCTATTTGTCAGGCTCATCGGTATGTTTCATCGTGTGATTTCAAGCAG 702  
Db 677 GCCTGCTGCTACTTTTTTCTGTGCTATTTGCAAGTTTCACTGTTGGAGATTTTA---AG 733  
Qy 703 CCTTACAAAGAGCAAGAGAGTTCAAGCTTTGGAGAGCAATGTT---CTATGTGGTGG 759  
Db 734 TGATAGCAAGAGAGCAAGAGAGTTCAAGATTGGAGATCAGTGTTTTACTATGCAATGA 793  
Qy 760 CTGTGTTTTCAGGCATCATATGCAAGGCTTCTTCTTGGAGGCAATGGAATTAATCTTCT 819  
Db 794 TAGTGATCACAGGAATAATATGCAAGGTTTCTTCTTAGGAGCCATAGGGATTTGTTT 853  
Qy 820 CCACATCGTCTCTCGTCTCGGTATTATGATCATGTCAGTGTGTTTCCCAATTTACAGAGTTT 879  
Db 854 GTGCATCATCACTAGCTTCTGGTGTCTGATAGTGTCTGCTTCGGTGCATGAAGTTT 913  
Qy 880 TAGCTGTTATTTCTACCATGAAAGTTTCAAGCTGAGAGGAGCTTCTCTTGTCTCTCT 939  
Db 914 TCGCGGTGCTTGTGTTTCCGGGAGAGTTTCAGGACAGAGAAAGGTGTCCTCTACTTCTT 973  
Qy 940 CCCTTTGGGGCTTTGTCCTTACTTTTATGGTGAGATAAAGTCTGCGGAGGA 991  
Db 974 CTCTTTGGGGAATTTGTCCTTACTTCTACGGGAGTTTAAATCCGGCAAGAA 1025

RESULT 3  
AAC44184  
ID AAC44184 standard; DNA; 2586 BP.  
XX  
AC AAC44184;  
XX  
DT 18-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 41938.

XX Hybridisation assay; genetic mapping; gene expression control;  
KW protein identification; signal transduction pathway; metabolic pathway;  
KW promoter; termination sequence; ss.  
XX Arabidopsis thaliana.  
PN EP1033405-A2.  
PD 06-SEP-2000.  
XX 25-FEB-2000; 2000EP-00301439.  
PF 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 18-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.

PR	23-JUN-1999;	99US-0140353P.	PR	31-AUG-1999;	99US-01514138P.
PR	23-JUN-1999;	99US-0140354P.	PR	01-SEP-1999;	99US-0151930P.
PR	24-JUN-1999;	99US-0140699P.	PR	07-SEP-1999;	99US-0152363P.
PR	28-JUN-1999;	99US-0140823P.	PR	10-SEP-1999;	99US-0153070P.
PR	29-JUN-1999;	99US-0140991P.	PR	13-SEP-1999;	99US-0153758P.
PR	30-JUN-1999;	99US-0141287P.	PR	15-SEP-1999;	99US-0154018P.
PR	01-JUL-1999;	99US-0141842P.	PR	16-SEP-1999;	99US-0154019P.
PR	01-JUL-1999;	99US-0142154P.	PR	20-SEP-1999;	99US-0154779P.
PR	02-JUL-1999;	99US-0142055P.	PR	22-SEP-1999;	99US-0155139P.
PR	06-JUL-1999;	99US-0142390P.	PR	24-SEP-1999;	99US-0155486P.
PR	08-JUL-1999;	99US-0142803P.	PR	28-SEP-1999;	99US-0155659P.
PR	09-JUL-1999;	99US-0142920P.	PR	29-SEP-1999;	99US-0156458P.
PR	12-JUL-1999;	99US-0142977P.	PR	04-OCT-1999;	99US-0156596P.
PR	13-JUL-1999;	99US-0143542P.	PR	04-OCT-1999;	99US-0157117P.
PR	14-JUL-1999;	99US-0143624P.	PR	05-OCT-1999;	99US-0157753P.
PR	15-JUL-1999;	99US-0144005P.	PR	06-OCT-1999;	99US-0157865P.
PR	16-JUL-1999;	99US-0144086P.	PR	07-OCT-1999;	99US-0158029P.
PR	16-JUL-1999;	99US-0144086P.	PR	08-OCT-1999;	99US-0158232P.
PR	19-JUL-1999;	99US-0144325P.	PR	12-OCT-1999;	99US-0158369P.
PR	19-JUL-1999;	99US-0144331P.	PR	13-OCT-1999;	99US-0159293P.
PR	19-JUL-1999;	99US-0144332P.	PR	13-OCT-1999;	99US-0159294P.
PR	19-JUL-1999;	99US-0144333P.	PR	14-OCT-1999;	99US-0159295P.
PR	19-JUL-1999;	99US-0144334P.	PR	14-OCT-1999;	99US-0159329P.
PR	19-JUL-1999;	99US-0144335P.	PR	14-OCT-1999;	99US-0159330P.
PR	20-JUL-1999;	99US-0144352P.	PR	14-OCT-1999;	99US-0159331P.
PR	20-JUL-1999;	99US-0144632P.	PR	14-OCT-1999;	99US-0159637P.
PR	20-JUL-1999;	99US-0144884P.	PR	14-OCT-1999;	99US-0159638P.
PR	21-JUL-1999;	99US-0144814P.	PR	18-OCT-1999;	99US-0159584P.
PR	21-JUL-1999;	99US-0145086P.	PR	21-OCT-1999;	99US-0160741P.
PR	21-JUL-1999;	99US-0145086P.	PR	21-OCT-1999;	99US-0160767P.
PR	22-JUL-1999;	99US-0145085P.	PR	21-OCT-1999;	99US-0160768P.
PR	22-JUL-1999;	99US-0145087P.	PR	21-OCT-1999;	99US-0160770P.
PR	22-JUL-1999;	99US-0145089P.	PR	21-OCT-1999;	99US-0160814P.
PR	22-JUL-1999;	99US-0145192P.	PR	21-OCT-1999;	99US-0160815P.
PR	23-JUL-1999;	99US-0145145P.	PR	22-OCT-1999;	99US-0160980P.
PR	23-JUL-1999;	99US-0145218P.	PR	22-OCT-1999;	99US-0160981P.
PR	23-JUL-1999;	99US-0145224P.	PR	22-OCT-1999;	99US-0160989P.
PR	26-JUL-1999;	99US-0145276P.	PR	25-OCT-1999;	99US-0161404P.
PR	27-JUL-1999;	99US-0145913P.	PR	25-OCT-1999;	99US-0161405P.
PR	27-JUL-1999;	99US-0145918P.	PR	25-OCT-1999;	99US-0161406P.
PR	27-JUL-1999;	99US-0145919P.	PR	26-OCT-1999;	99US-0161359P.
PR	28-JUL-1999;	99US-0145951P.	PR	26-OCT-1999;	99US-0161360P.
PR	02-AUG-1999;	99US-0146386P.	PR	26-OCT-1999;	99US-0161361P.
PR	02-AUG-1999;	99US-0146388P.	PR	28-OCT-1999;	99US-0161920P.
PR	02-AUG-1999;	99US-0146389P.	PR	28-OCT-1999;	99US-0161922P.
PR	03-AUG-1999;	99US-0147038P.	PR	28-OCT-1999;	99US-0161993P.
PR	04-AUG-1999;	99US-0147204P.	PR	29-OCT-1999;	99US-0162142P.
PR	04-AUG-1999;	99US-0147302P.	Query Match 32.1%; Score 342.8; DB 3; Length 2586;		
PR	05-AUG-1999;	99US-0147192P.	Best Local Similarity 67.9%; Pred. No. 8.6e-89;		
PR	05-AUG-1999;	99US-0147260P.	Matches 509; Conservative 1; Mismatches 234; Indels 6; Gaps 2;		
PR	06-AUG-1999;	99US-0147303P.			
PR	06-AUG-1999;	99US-0147416P.	Qy	245	TCTTATCGCGCTGTTATTGTAGGCATTCTCTCAGGCTTCTTATCACTTCTTACAGTTAGCTTT
PR	09-AUG-1999;	99US-0147493P.	Db	1643	TCATATCGCTCCATTGTCATAGGTTGCTCAGAGGACTTGACAACTACTTATATTCTTA
PR	09-AUG-1999;	99US-0147935P.	Qy	305	TGGTATAGCTTATCTTCCAGTTTCTCAGCTGCTCTTATCATTCCTTACAGTTAGCTTT
PR	10-AUG-1999;	99US-0148171P.	Db	1703	CGGATTAGCATATCTGCCAGTTTCACTTCATCGCTCATATCGGAACTCAACTAGCTTT
PR	11-AUG-1999;	99US-0148319P.	Qy	365	TATAGCTATCTTCTCATCTTTCATGGTTAAACATAAAGTTCACTCTCTTTTACCATCAATGC
PR	12-AUG-1999;	99US-0148341P.	Db	1763	CAACGCTCTCTTGGCTTCTTGTAGTCAAGCAAAAGTTTCACTCCGTTCTCCATAAAGC
PR	13-AUG-1999;	99US-0148565P.	Qy	425	TGTTGCTTGTGACTGTTGGTGCTCGGTTTTGGGAATGCATACCGAAACTGATAAGCC
PR	16-AUG-1999;	99US-0149368P.	Db	1823	CGTCGTTTTTGTGACGGTTGGTATCGGGATCCCTTCGCTTACACAGTGCATGGAGACAACC
PR	17-AUG-1999;	99US-0149175P.	Qy	485	AGTTCATGAGACTCACAAAGCAGTACATAAAGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTT
PR	18-AUG-1999;	99US-0149426P.	Db	1882	CGTCTGTTTGTGACGGTTGGTATCGGGATCCCTTCGCTTACACAGTGCATGGAGACAACC
PR	20-AUG-1999;	99US-0149722P.	Qy	544	AGTTCATGAGACTCACAAAGCAGTACATAAAGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTT
PR	20-AUG-1999;	99US-0149722P.	Db	1942	GGCTAAGGAGAGCAAGAAAGAGTATGTTGGGTTCTTCTTCTTCTTCTTCTTCTTCTTCTT
PR	23-AUG-1999;	99US-0149902P.	Qy	545	TATGATGCTTTCATCTTGCCATTAGTGGAACCTTGCTTACCAGAAAGCTAAGCAACCAT
PR	23-AUG-1999;	99US-0149930P.	Db		
PR	25-AUG-1999;	99US-0150566P.	Qy		
PR	26-AUG-1999;	99US-0150884P.	Db		
PR	27-AUG-1999;	99US-0151065P.	Qy		
PR	27-AUG-1999;	99US-0151066P.	Db		
PR	27-AUG-1999;	99US-0151080P.	Qy		
PR	30-AUG-1999;	99US-0151303P.	Db		

Db	1943	TCTCTATGCTTTTATATATACCGTCTGTGAGCTAACTTACAAGAAAGCTCGTCAAGAAAT	2002	PR	06-MAY-1999;	99US-0132486P.
QY	605	GAGCTATACCTTGTCCTGAGTTCCAGTTGATTTTGTGTCTCTCTTCTTCTATTCTCAG	664	PR	06-MAY-1999;	99US-0132487P.
Db	2003	CACCTTCCCACTTGCTTGAGATTGAGATGTCATGCTTGCTGCTACTTTTTTCG	2062	PR	11-MAY-1999;	99US-0132863P.
QY	665	CGTCATCGGTATGTTTCATCGCTGGTGATTTCAAGCAGGCGCTTACCAAAAGAACGAAGAGA	724	PR	11-MAY-1999;	99US-0134256P.
Db	2063	TGTCATTTGGCATGTTTCATCGTTGGAGATTTTA---AGGTGATGCAVGAAGCAAGAGA	2119	PR	14-MAY-1999;	99US-0134218P.
QY	725	GTTCAAGCTTGGAGAGCATGTT---CTATGGTGGCTGTGTTTTTCAGCCATCATATG	781	PR	14-MAY-1999;	99US-0134219P.
Db	2120	GTTCAAGATTGGAGGATCAGTGTTTTACTATGCAATTGATGTCATCAGGAATAATATG	2179	PR	14-MAY-1999;	99US-0134221P.
QY	782	GCAAGGCTTCTCTGGAGCCATGGAATTAATCTTCTCCACATCGTCTCTGCTCGGG	841	PR	18-MAY-1999;	99US-0134370P.
Db	2180	GCAAGTTTCTCTTAGGAGCCATAGGATGTGTTTTGTGCATCATCACTAGCTTCTGG	2239	PR	18-MAY-1999;	99US-0134768P.
QY	842	TATTATGATATCAGTCTCTTTTGCCAAATTACAGAGGTTTTTAGCTGTGTTATATTTACCATGA	901	PR	20-MAY-1999;	99US-0134941P.
Db	2240	TGTTCTGATAAGTGTCTGCTTCCGGTGACTGAAGTTTTTCGCGTCTGTTTTCGGGA	2299	PR	20-MAY-1999;	99US-0135124P.
QY	902	AAAGTTTCAAGCTGAGAGGACCTTCTCTGCTCTCTCCCTTTGGGGCTTTGTCTCTTA	961	PR	21-MAY-1999;	99US-0135353P.
Db	2300	GAAGTTTTCAGGCAGAGAAAGTGTCCTCTACTTCTTCTCTTTGGGGATTTGTCTCTTA	2359	PR	24-MAY-1999;	99US-0135629P.
QY	962	CTTTTATGTGACATAAAGTCTCGCGAGGA	991	PR	25-MAY-1999;	99US-0136021P.
Db	2360	CTTCTACGGCGAGTTTAAATCCGGCAAGAA	2389	PR	27-MAY-1999;	99US-0136392P.
RESULT 4						
AAC36831						
ID	AAC36831 standard; DNA; 493 BP.					
XX	AC AAC36831;					
XX	AC AAC36831;					
DT	17-OCT-2000 (first entry)					
DE	Arabidopsis thaliana DNA fragment SEQ ID NO: 15222.					
KW	Hybridisation assay; genetic mapping; gene expression control;					
KW	protein identification; signal transduction pathway; metabolic pathway;					
KW	promoter; termination sequence; ss.					
OS	Arabidopsis thaliana.					
XX	EP1033405-A2.					
XX	EP1033405-A2.					
PD	06-SEP-2000.					
PF	25-FEB-2000; 2000EP-00301439.					
XX	25-FEB-1999; 99US-0121825P.					
PR	05-MAR-1999; 99US-0123180P.					
PR	09-MAR-1999; 99US-0123548P.					
PR	23-MAR-1999; 99US-0125788P.					
PR	25-MAR-1999; 99US-0126264P.					
PR	29-MAR-1999; 99US-0126785P.					
PR	01-APR-1999; 99US-0127462P.					
PR	06-APR-1999; 99US-0128234P.					
PR	08-APR-1999; 99US-0128714P.					
PR	16-APR-1999; 99US-0129845P.					
PR	19-APR-1999; 99US-0130077P.					
PR	21-APR-1999; 99US-0130449P.					
PR	23-APR-1999; 99US-0130510P.					
PR	28-APR-1999; 99US-0130891P.					
PR	30-APR-1999; 99US-0131449P.					
PR	30-APR-1999; 99US-0132048P.					
PR	04-MAY-1999; 99US-0132407P.					
PR	05-MAY-1999; 99US-0132484P.					
PR	05-MAY-1999; 99US-0132485P.					
PR	06-MAY-1999; 99US-0132486P.					
PR	07-MAY-1999; 99US-0132487P.					
PR	11-MAY-1999; 99US-0132863P.					
PR	14-MAY-1999; 99US-0134218P.					
PR	14-MAY-1999; 99US-0134219P.					
PR	14-MAY-1999; 99US-0134221P.					
PR	18-MAY-1999; 99US-0134370P.					
PR	19-MAY-1999; 99US-0134941P.					
PR	20-MAY-1999; 99US-0135124P.					
PR	21-MAY-1999; 99US-0135353P.					
PR	24-MAY-1999; 99US-0135629P.					
PR	25-MAY-1999; 99US-0136021P.					
PR	27-MAY-1999; 99US-0136392P.					
PR	28-MAY-1999; 99US-0136782P.					
PR	01-JUN-1999; 99US-0137222P.					
PR	03-JUN-1999; 99US-0137528P.					
PR	04-JUN-1999; 99US-0137502P.					
PR	07-JUN-1999; 99US-0137724P.					
PR	08-JUN-1999; 99US-0138094P.					
PR	10-JUN-1999; 99US-0138540P.					
PR	10-JUN-1999; 99US-0138847P.					
PR	14-JUN-1999; 99US-0139119P.					
PR	16-JUN-1999; 99US-0139452P.					
PR	16-JUN-1999; 99US-0139453P.					
PR	17-JUN-1999; 99US-0139492P.					
PR	18-JUN-1999; 99US-0139454P.					
PR	18-JUN-1999; 99US-0139455P.					
PR	18-JUN-1999; 99US-0139456P.					
PR	18-JUN-1999; 99US-0139457P.					
PR	18-JUN-1999; 99US-0139458P.					
PR	18-JUN-1999; 99US-0139459P.					
PR	18-JUN-1999; 99US-0139460P.					
PR	18-JUN-1999; 99US-0139461P.					
PR	18-JUN-1999; 99US-0139462P.					
PR	18-JUN-1999; 99US-0139463P.					
PR	18-JUN-1999; 99US-0139750P.					
PR	18-JUN-1999; 99US-0139763P.					
PR	21-JUN-1999; 99US-0139817P.					
PR	22-JUN-1999; 99US-0139899P.					
PR	23-JUN-1999; 99US-0140353P.					
PR	23-JUN-1999; 99US-0140354P.					
PR	24-JUN-1999; 99US-0140695P.					
PR	28-JUN-1999; 99US-0140823P.					
PR	29-JUN-1999; 99US-0140991P.					
PR	30-JUN-1999; 99US-0141287P.					
PR	01-JUL-1999; 99US-0141842P.					
PR	01-JUL-1999; 99US-0142154P.					
PR	02-JUL-1999; 99US-0142055P.					
PR	06-JUL-1999; 99US-0142390P.					
PR	08-JUL-1999; 99US-0142803P.					
PR	09-JUL-1999; 99US-0142920P.					
PR	12-JUL-1999; 99US-0142977P.					
PR	13-JUL-1999; 99US-0143542P.					
PR	14-JUL-1999; 99US-0143624P.					
PR	15-JUL-1999; 99US-0144005P.					
PR	16-JUL-1999; 99US-0144085P.					
PR	16-JUL-1999; 99US-0144086P.					
PR	19-JUL-1999; 99US-0144325P.					
PR	19-JUL-1999; 99US-0144331P.					
PR	19-JUL-1999; 99US-0144332P.					
PR	19-JUL-1999; 99US-0144333P.					
PR	19-JUL-1999; 99US-0144334P.					
PR	19-JUL-1999; 99US-0144335P.					
PR	20-JUL-1999; 99US-0144352P.					
PR	20-JUL-1999; 99US-0144632P.					
PR	20-JUL-1999; 99US-0144684P.					
PR	21-JUL-1999; 99US-0144814P.					
PR	21-JUL-1999; 99US-0145086P.					
PR	21-JUL-1999; 99US-0145088P.					
PR	22-JUL-1999; 99US-0145085P.					
PR	22-JUL-1999; 99US-0145087P.					











Db 847 GAGACCTTTCTCAGTTGTGTGAACATGCAGATATATACAGCTCTCGTGGCAACATTGGCT 906  
QY 663 AGCGTCATCGGTATGTTTCATCGCTGGTGAATTCAG 698  
Db 907 TCTCTTGTGGGTATATTGGCAAGTGGTGAATGGATG 942

RESULT 9  
AAA97923  
ID AAA97923 standard; DNA; 1194 BP.  
XX AC AAA97923;  
XX 19-JAN-2001 (first entry)  
XX DE A. thaliana PUP1 DNA #5.  
XX KW PUP1; transgenic plant; nucleobase transporter; apical dominance;  
XX KW flowering behaviour; senescence; pesticide distribution; ds.  
XX OS Arabidopsis thaliana.  
XX DE19907209-A1.  
XX PN 19-FEB-1999; 99DE-01007209.  
XX PD 19-FEB-1999; 99DE-01007209.  
XX PA (FROM/) FROMMER W.  
XX PI Gillissen B, Buerkle L, Andre B, Frommer WB;  
XX DR WPI; 2000-566202/53.  
XX PT Nucleic acid, useful for producing transgenic plants with altered  
PT nucleobase transport, encodes a nucleobase transporter protein of  
PT Arabidopsis thaliana.  
XX PS Claim 1f; Page 14; 24pp; German.

XX This invention describes a novel nucleic acid encoding a plant nucleobase  
CC transporter (I). (I) is produced by complementation of a nucleobase  
CC transport (NBT)-defective host cell with a plant gene bank by selection  
CC of NBT-positive cells. (I) is used to isolate homologous sequences from  
CC bacteria, fungi, plants, animals and humans, for expression of the  
CC encoded protein (II) in prokaryotic or eukaryotic cells, for inhibiting  
CC expression of (II) (when in antisense orientation), and to produce  
CC transgenic crop plants. The transgenic plants have modified nucleobase  
CC transport properties, e.g. altered affinity and substrate specificity  
CC that may result in more efficient nucleobase transport in leaves, changes  
CC in apical dominance, flowering behaviour and senescence, or improved  
CC distribution of pesticides. This sequence encodes the Arabidopsis  
CC thaliana PUP1 protein which is described in the method of the invention

XX Sequence 1194 BP; 298 A; 279 C; 231 G; 386 T; 0 U; 0 Other;  
SQ

Query Match 12.8%; Score 136.2; DB 3; Length 1194;  
Best Local Similarity 48.7%; Pred. No. 8.4e-29;  
Matches 462; Conservative 0; Mismatches 478; Indels 9; Gaps 3;

QY 20 TCTTGTGATCAATAACTGCATAAATCTAGCCATAGGAAACTGTGGAGTCCTTTGATAT 79  
Db 135 TCGTGTCTCCATATACGPAATCTTTGTCTCTTTTGGCAGCCACTAGTACAAATCTGGG 194  
QY 80 CGGTCTCTACTTCAACAATGCGGGTAAAGGATTTGGTCTCTAGCTTTCTTGAACATGC 139  
Db 195 TAGATTGACTATGAANAATGAGGNAATAGACATATGTGGTAACACTCTTCACTCAT 254  
QY 140 AGGCTTTCCTGTATCTTTCATCTCCTCTCTCTTTCATATTACCCGGCCGAGAACAA 199  
Db 255 TGGCTTCCCTACTGTTCTGTCTCGGCTTCTTTTCTCGAATCAGGCAACCCAAATCAAC 314

QY 200 CAATGTGGTGATAGTACAAAGTTTCTTTCTTATCAAAACCGCGTCTTCTTATCGCGGTGT 259  
Db 315 AGATACAAATTTTCAGT--CAGTCCCCTTCTCTTACCACCCCTTGCATCGGTACTTGTGC 372  
QY 260 TATTGTAGGCAATCTCTCAGGGTTTGTAACTACTTCTGTATGCAATATGATAGCTTATCT 319  
Db 373 ACT---GGACTGCTAGTGTCCGCTTATGCTTATTGTCTGCAGTAGGGTGTCTACTT 428  
QY 320 TCCAGTTTCTACAGCTGCTCTTATCATTTGTTTCTACAGTTAGCTTTTATAGCTATCTTCTC 379  
Db 429 ACCAGTCTCTACTTTTCTCCCTCATCTTGGCTCAGAGTTGSCCTTCTACTGCTTTTCTC 488  
QY 380 ATTCTTTCATGTTAAACATAAAGTTCACTCTCTTTTACCATCAATGCTCTGTTGTTGAC 439  
Db 489 ATATTTCCTTAACTCGCAAAAGTTCACTCTCTTTGATAGTCAGTTCTTTTGTCTTCCTCAC 548  
QY 440 TGTGTGTGCTGCGGTTTTTGGGAATGCATACGAAACTGATAAGCAGTTTATGAGACTCA 499  
Db 549 TGTATCTCTGCTCTTCTTCTTGTGTCAACACTGATTGAGAAAACTCAACTAATGTATCTAG 608  
QY 500 CAAGCAGTACATAAATCGTTTCTTGTATTACTGTAGCAGCAGCTGTTATGATGTTTCAAT 559  
Db 609 AGTACAGTATGTATCGGTTTCATATGATACCATCGGTGCTTCCGCTGGGATTGACTGTT 668  
QY 560 CTTGCCATTAGTGGAACTTGTCTTACCAGAAAGCTTAAGCAACCACTAGACTATACCTTGT 619  
Db 669 ACTATCTCTGATACAAATGCTCTTTCAGGAAAGTTTTTTCAGGAAGCATACATCCTCAGCAGT 728  
QY 620 GCTCGAGTTCCAGTTGATTTTGTGTCTCTTGTCTTCTTATTGTCTCAGCGTCATCGGTATGTT 679  
Db 729 CACGGAATTTGGCCATTTTACCAGTCTCTAGTTGCGAGTTGTAGTTCTCATAGGACTTTT 788  
QY 680 CATCGCTGGTGATTTCAAGCAGCGCTTACCAAAAGAGCAAGAGAGTTCAGAGTTGGAGA 739  
Db 789 TGCAAGTGGAGAGT---GGGAAACTTTGCCAAGTGAGATGAGAACTACAAACTCGGGAA 845  
QY 740 GGCAATTTCTATGTGTGTGCTGTGTTTTCAGCCATCATATGCAAGGCTCTTCTTCTGGG 799  
Db 846 AGTGTCTATATGTTTGTAGCTTTCAGCTCGGAGCTATTTCTCTGCAAGTCTACACTTTGG 905  
QY 800 AGCCATTGGATTAATCTTCTCCACATCGTCTCTCGGTCGCGGTATTATGATATCAGTGTCT 859  
Db 906 TCTTGTGGGATGATCTTCGAGTCACTCTGTGTCTCTCCAAATTCATACAGCTGTGGG 965  
QY 860 TTTGCCAATTACAGAGGTTTGTAGCTGTATATTTCTACCATGAAAGTTTCAAGCTGAGAA 919  
Db 966 ATTGCTTATAGTTCCAGTTTGGGCGAGTGATAGTTTTCATGATAGAAATGGACGCATCAAA 1025  
QY 920 GGGACTTTCTCTGCTCTCTCCCTTTGGGGCTTTGTCTCTTACTTTTAT 968  
Db 1026 AATCTTCTCCATTATTTAGCTATCTCGGGCTTCTCTTCAATTCGTCTAT 1074

## RESULT 10

ADG88190  
ID ADG88190 standard; cDNA; 1047 BP.

XX AC ADG88190;

XX AC ADG88190;

DT 22-APR-2004 (first entry)

XX AC ADG88190;

DE A. thaliana RPP4-upregulated pathogen infection-related gene #632.

XX Pathogen infection-related gene; plant; Peronospora parasitica;

XX defence mechanism; RPP4; pathogen resistance; transgenic plant; oomycete;

XX fungus; bacterium; virus; nematode; insect; aphid; gene; ss.

OS Arabidopsis thaliana.

XX WO200222675-A2.

PD 21-MAR-2002.



PT Arabidopsis thaliana.  
XX Claim 2; Page 14-15; 24pp; German.  
XX This invention describes a novel nucleic acid encoding a plant nucleobase  
CC transporter (I). (I) is produced by complementation of a nucleobase  
CC transport (NBT)-defective host cell with a plant gene bank by selection  
CC of NBT-positive cells. (I) is used to isolate homologous sequences from  
CC bacteria, fungi, plants, animals and humans, for expression of the  
CC encoded protein (II) in prokaryotic or eukaryotic cells, for inhibiting  
CC expression of (I) (when in antisense orientation), and to produce  
CC transgenic crop plants. The transgenic plants have modified nucleobase  
CC transport properties, e.g. altered affinity and substrate specificity  
CC that may result in more efficient nucleobase transport in leaves, changes  
CC in apical dominance, flowering behaviour and senescence, or improved  
CC distribution of pesticides. This sequence encodes the Arabidopsis  
CC thaliana PUP1 protein which is described in the method of the invention  
XX  
SQ Sequence 1081 BP; 322 A; 231 C; 205 G; 323 T; 0 U; 0 Other;  
Query Match 12.7%; Score 135; DB 3; Length 1081;  
Best Local Similarity 49.6%; Pred. No. 1.8e-28;  
Matches 438; Conservative 0; Mismatches 430; Indels 15; Gaps 3;  
QY 85 TCTACTTCAACAATGGCGGTAAGGATTTGGTCTCTACGTTTCTTGAAACTGCAGGCT 144  
Db TATACTATGACAAACGAGGAAACAGTAATGGTAGCAACGGTAGTTTCAACTTGTGGCT 151  
QY 145 TTCCTGTTATCTTCAATTCCTCGCTCTCTTCTTCTTACATTAACCGGGCGAGAACAAATG 204  
Db TTTCTGTGCTACTTCCATATTATATCTTGTTCATTTAAACACATGCAACA-----ACTG 205  
QY 205 TGGGTGATGATACAGATTTCTTCTTATCAACCGGCTTCTTATCGCCGCTGTTATTG 264  
Db ATAGAGATGGAAAGAACCTCACCTAGGAA-----CCGTGATTTGGTTACGTAGTGC 259  
QY 265 TAGGCATTTCTCTCAGGGTTTGATACACTACTTGTATGCATATGGTATAGCTTATCTTCAG 324  
Db TTGGACTTCTTGTAGGACAGATTGCTATCTGTACTCCATTGGACTTCTTTACTTACCG 319  
QY 325 TTTCTACAGCTGCTTTATCATTTGTTTACAGTTAGCTTTTATAGCTATCTTCTCATTTCT 384  
Db TTTCTACCTATTCCTCGTATCTGTGCATCTCAGTTAGCCTTCAATGCTTCTTCTCTATT 379  
QY 385 TCATGGTTAAACATAAAGTTCACTCTCTTTTACCATCAATGCTGTGTGTGTTGACTGTG 444  
Db TTTCTAACTCACAAAACCTACCCCTATCATTTTAAATCTCTTTCTCTCTACTATAT 439  
QY 445 GTGCTGCGGTTTTGGGAATGCATACCGAAACTGATAGCCAGTTTCATGAGACTCACAGC 504  
Db CTTCACCCCTACTTGCATTCAATTAATAGGAGACAGACTCCACAAAAGTTACAAAAGGAG 499  
QY 505 AGTACATAACTGGTTTCTTGATTACTGTAGCAGAGCTGTTATGTATGCTTTCATCTTGC 564  
Db AGTATGTCAAAAGGTTTCATATGCACCGTTGCTGGCTCTGCTGGTTATGGTCTAGTCTTAT 559  
QY 565 CATTAGTGGAACTTCTTACCAGAAAGCTAAGCAAAACCATGAGCTATACCCCTTGTGCTCG 624  
Db CCCTACACAGCTAGCCCTTTCTAAAGTCCTTAAGAACCAAAATTTCTCAAGATTATGG 619  
QY 625 AGTTCCAGTTGATTTGTGTCTCTCTTGTCTTATTTGTCAGCGTTCATCGTATGTTTCATCG 684  
Db ATATGATAATCTACGTAGTCTTAGTGGCCAGTTGTGTTAGCGTGTGGGGCTTT--TTG 676  
QY 685 CTGGTATTTCAAGCAGCCCTTACAAAGAACGACAGAGTTTCAGCTTTCGAGGAGCAT 744  
Db CTACAGTGTAGTGGAAAACTTTTGACAGTGAATGGATAAATACAAAACATGGGAAGGTAT 736  
QY 745 TGTTCATGTGTGTGCTGTGTTTTTCAGGCATCATATGGCAAGGCTTCTTCTTTGGAGCCA 804  
Db CCTACATATTGAACCTAGTGTGGACAGCTGTATCTTGGCAGGTATTTCTCCATCGTGGCA 796  
QY 805 TTGGATTAACTTTCTCCACATCGTCTCTCGTCTCGGGTATTATGATATCATGCTGCTTTTGC 864

Db 797 CAGGACTGATCTTCGAGCTCTCTCTCTATTTCTCAAAATGCAATAAGCGTTTTCGGACTCC 856  
QY 865 CAATTACAGAGGTTTACGCTGTTATATCTACCATGAAGAAGTTTCAAGCTGAGAAGGAC 924  
Db 857 CAGTGGTTCCTATCTTGGCTGTAATCATTTTCCATGACAAAATGAATGGGTTAAAGGTGA 916  
QY 925 TTTCTCTGTCTCTCTCCCTTTGGGCTTTGTCTCTTACTTTTA 967  
Db 917 TTTCTATGATCCTAGCTATTGGGTTTCACTTCTTATGCTA 959  
RESULT 12  
ADA70627  
ID ADA70627 standard; DNA; 1200 BP.  
XX ADA70627;  
AC ADA70627;  
XX 20-NOV-2003 (first entry)  
XX Rice gene, SEQ ID 3950.  
XX Plant; bacterial infection; fungal infection; viral infection; rice;  
XX gene; ds.  
XX Oryza sativa.  
XX WO2003000898-A1.  
XX 03-JAN-2003.  
XX 22-JUN-2001; 2001WO-IB001105.  
XX 22-JUN-2001; 2001WO-IB001105.  
XX (SYGN ) SYNGENTA PARTICIPATIONS AG.  
XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;  
XX Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;  
XX WPI; 2003-175290/17.  
XX Identifying at least one gene involved in plant resistance or response to  
XX pathogenic infection for conferring resistance or tolerance to a plant to  
XX bacterial, fungal or viral infection by determining or detecting plant  
XX gene expression.  
XX Claim 6; SEQ ID NO 3950; 899pp; English.  
XX The present invention relates to a method (M1) for identifying genes  
XX involved in plant resistance or response to pathogenic infection. M1  
XX comprises identifying a gene whose expression is significantly altered in  
XX the incompatible interaction of plant gene expression relative to  
XX expression of the gene in an uninfected plant, in a mutant plant that  
XX does not express a gene associated with response to pathogenic infection,  
XX or in a corresponding incompatible or compatible interaction. (M1) is  
XX useful for conferring resistance to resistance or tolerance to a plant to  
XX bacterial, fungal or viral infection. The present sequence was used to  
XX illustrate the invention.  
SQ Sequence 1200 BP; 143 A; 416 C; 434 G; 206 T; 0 U; 1 Other;  
Query Match 12.3%; Score 131.4; DB 8; Length 1200;  
Best Local Similarity 48.8%; Pred. No. 2.1e-27;  
Matches 480; Conservative 0; Mismatches 461; Indels 42; Gaps 3;  
QY 17 GGCTCTTGTGATCAATAAATGCATAATTTAGCCATAGGAAACTGTGAGGTCTTTGAT 76  
Db 138 GTCGCTGCTGATATTCAGCGCTGCTCTCTCTCGCGCGCGCGCGCTCT 197  
QY 77 TATCGCTCTCTACTTTCACAAATGGCGGTAAAGATTGGTTCTCTCTACGTTCTTTGAAC 136  
Db 198 CCTCCGCGTCTACTTCTGTCACGCGGGACCCCGCTGTGCTGTCTCGCCAGGCTCCAGAT 257



Db 707 TAAATATGTATAGGATATATATGACATTTGGTCTCTCTGGGATTTGGATGCTGC 766  
Qy 562 TGGCAATATAGTGAATCTCTACCAAGAAAGTAAAGCAACCATGAGCTATACCTTTGTC 621  
Db 767 TATCCCTGGTACAACTGATCCTCAGGAAGGTTTAAAGAACCAATCTCTCAACGGTCA 826  
Qy 622 TCGAGTTCAGTATTTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 681  
Db 827 CTGACTTGGTCTGCTTACCAATCTCTAGTTGCAAGCTGTGGTCTCATAGGACTTTTCG 886  
Qy 682 TCGTGTGTGATTTCAAGCAGGCTTTACCAAGAAAGCAAGAGATTCAAGCTTTGGAGAG 741  
Db 887 CAACGGGGAGT--GGAAACTTTAACTGAGATGGAATACTACAACTGGGAAAG 943  
Qy 742 CATTTGTTCTATGTGTGCTGTGTTTTCAGGCATCATATGCAAGGCTTCTCTTTGGAG 801  
Db 944 TGCCATAGCTTATGACTTTGGCTCGATAGCTATTTCTCTGGCAAGTCTACACCATTTGGC 1003  
Qy 802 CCATTGGATTAATCTTCCACATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 861  
Db 1004 TCGTGGGACTGATCTTTGAGTCATCTCTGTGTTCTCCAAATTCATAACTCTGTGGAT 1063  
Qy 862 TGCCAAATPACAGAGGTTTTAGCTGTATATTTCTACCATGAAAGTTTCAAGCTCGAAGG 921  
Db 1064 TGCCTATAGTTCAGTGTAGCAGTATGTTTCCATGATAAATGAACGGTCAAGA 1123  
Qy 922 GACTTTCTCTGCT 968  
Db 1124 TCTTCTCCATCATTTTAGCTATCTGGGATTCATTTCAATTTGTTCTAT 1170

## RESULT 14

AAA97921  
ID AAA97921 standard; DNA; 1145 BP.

XX AC AAA97921;

XX XX 19-JAN-2001 (first entry)

XX DE A. thaliana PUP1 DNA #3.

XX PUPL; transgenic plant; nucleobase transporter; apical dominance;  
KW flowering behaviour; senescence; pesticide distribution; ds.

XX OS Arabidopsis thaliana.

XX PN DE19907209-A1.

XX PD 24-AUG-2000.

XX PF 19-FEB-1999; 99DE-01007209.

XX PR 19-FEB-1999; 99DE-01007209.

XX PA (FROM/) FROMMER W.

XX PI Gillissen B, Buerkle L, Andre B, Frommer WB;

XX DR WPI; 2000-566202/53.

PT Nucleic acid, useful for producing transgenic plants with altered  
PT nucleobase transport, encodes a nucleobase transporter protein of  
XX Arabidopsis thaliana.

XX PS Claim 1f; Page 12-13; 24pp; German.

XX CC This invention describes a novel nucleic acid encoding a plant nucleobase  
CC transporter (I). (I) is produced by complementation of a nucleobase  
CC transport (NBT)-defective host cell with a plant gene bank by selection  
CC of NBT-positive cells. (I) is used to isolate homologous sequences from  
CC bacteria, fungi, plants, animals and humans, for expression of the  
CC encoded protein (II) in prokaryotic or eukaryotic cells, for inhibiting

CC expression of (II) (when in antisense orientation), and to produce  
CC transgenic crop plants. The transgenic plants have modified nucleobase  
CC transport properties, e.g. altered affinity and substrate specificity  
CC that may result in more efficient nucleobase transport in leaves, changes  
CC in apical dominance, flowering behaviour and senescence, or improved  
CC distribution of pesticides. This sequence encodes the Arabidopsis  
CC thaliana PUP1 protein which is described in the method of the invention

XX SQ Sequence 1145 BP; 309 A; 234 C; 226 G; 376 T; 0 U; 0 Other;

Query Match 10.9%; Score 116.6; DB 3; Length 1145;  
Best Local Similarity 48.3%; Pred. No. 4.2e-23;  
Matches 434; Conservative 0; Mismatches 444; Indels 21; Gaps 3;

Qy 82 GTCTCTACTTCAACAATGGCGGTAAAGAGATTTGGTCTCTACGTTTCTTGAAGACTGCAG 141  
Db 140 GACTTTACTAGAAAAAGCGGTAAAGACATGGCTCGAAACCTTGGTTTCAGCTTTAG 199  
Qy 142 GCTTTCTGTTATCTTCAATTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 201  
Db 200 GGTTCCTTTTAAACCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 259  
Qy 202 ATGTGGGTGATAGTACAAAGTTTCTTTCTTATCAAAACCGCTCTTCTTATCGCGCTGTA 261  
Db 260 CCATTACCAAAAAAATCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 313  
Qy 262 TTGTAGGCATCTCTCAGGTTTGTATTAACACTCTCTCTCTCTCTCTCTCTCTCTCTCT 321  
Db 314 GACTTGGCTTGTCTTGTCTGGAATTTGTACTCATTTGGGCTACTTTTACCTTC 373  
Qy 322 CAGTTTCTACAGCTGCTCTTATCATCTCTTACAGTTTGTATTTTATAGCTATCTTCTCAT 381  
Db 374 CTGTCTCACTTCT 433  
Qy 382 TCTTCATGTTTAAACATAAAGTTTCACTCTTTTACCATCAATGCTGTGTGTGTGACTG 441  
Db 434 ACTTCTCTAAACTCACAAAAAATCAGACCAATTTTACTCAATTCACTTGTTCTCTTAA 493  
Qy 442 TTGGTCTGCGGTTTGGGAAATGCATACCGAAACTGA-----TAAGCCAGTTC 489  
Db 494 TATCTCTACACTTCTTGTATCCACATGAACAGAAATCTCCCTCTTCTTACTTCAAGT 553  
Qy 490 ATGAGACTCAACAGCAGTACATAAAGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 549  
Db 554 CCGCAGCAAGTCCAAGTATGTGATACATCTCGCGGTCTGCTAGCTAGCTAGCTGTT 613  
Qy 550 ATGCTTTTCATCTTGCCTATAGTGAACCTTCTTACCAGAAAGCTTAAGCAACCAAGCT 609  
Db 614 ATCTCTGGTGTCTTCTTAAACAGATTACGCGTTTCGAAAGATTCTTAAAGAAATACACAT 673  
Qy 610 ATACCTTGTCTCGAGTTCCAGTTGATTTTGTGTCTCTCTCTCTCTCTCTCTCTCTCTCT 669  
Db 674 TCAAGGCTATTTTAGACATGGCCACATATCCGCTCTATGGTAGCTACTTGTGTAGTTGG 733  
Qy 670 TCGGTATGTTTCATCGCTGGTGAATTTCAAGCAGGCTTTACCAAGAAAGCAAGAGATTCA 729  
Db 734 TAGGACTTTTGGAGTGGTGGTGGGAAAG---CTGAGTACAGAAATGGAAGATTTC 790  
Qy 730 AGCTTGGAGAGGCAATTTCTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 789  
Db 791 AACTTAGGAAAAAGCTCATACATTTTGATAAAACATCGGTTCAACGATATCATATGGCAAGCTT 850  
Qy 790 TCTTCTTGGGAGCCATTTGGATTAATCTTCTCCACATCGTCTCTCTCTCTCTCTCTCTCTCT 849  
Db 851 GTTTGATTGGAAGT 910  
Qy 850 TATCAGTGTCTTCTCCAAATTACAGAGTTTGTAGCTTTATATCTTACCATGAAAGTTTC 909  
Db 911 GCACTCTTGTTTTACCAAGTTGTGCTGTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 970  
Qy 910 AAGCTGAGAGAGGACTTTCT 968  
Db 971 GTGGAATCAAGTTGGTTGCAATGTTTTTGGCCATCTGGGATTTGTTTCTTATGGTTAT 1029



```
RESULT 15
AAC51119
ID AAC51119 standard; DNA; 1292 BP.
XX
AC AAC51119;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 67344.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD
PF
PF
XX
XX 25-FEB-2000; 2000EP-00301439.
XX
XX 25-FEB-1999; 99US-0121825P.
XX 05-MAR-1999; 99US-0123180P.
XX 09-MAR-1999; 99US-0123548P.
XX 23-MAR-1999; 99US-0125788P.
XX 25-MAR-1999; 99US-0126264P.
XX 29-MAR-1999; 99US-0126785P.
XX 01-APR-1999; 99US-0127462P.
XX 06-APR-1999; 99US-0128234P.
XX 08-APR-1999; 99US-0128714P.
XX 16-APR-1999; 99US-0129845P.
XX 19-APR-1999; 99US-0130077P.
XX 21-APR-1999; 99US-0130449P.
XX 23-APR-1999; 99US-0130510P.
XX 28-APR-1999; 99US-0130891P.
XX 30-APR-1999; 99US-0131449P.
XX 30-APR-1999; 99US-0132048P.
XX 04-MAY-1999; 99US-0132407P.
XX 05-MAY-1999; 99US-0132484P.
XX 06-MAY-1999; 99US-0132485P.
XX 06-MAY-1999; 99US-0132486P.
XX 07-MAY-1999; 99US-0132487P.
XX 11-MAY-1999; 99US-0132863P.
XX 14-MAY-1999; 99US-0134256P.
XX 14-MAY-1999; 99US-0134218P.
XX 14-MAY-1999; 99US-0134219P.
XX 14-MAY-1999; 99US-0134221P.
XX 18-MAY-1999; 99US-0134370P.
XX 19-MAY-1999; 99US-0134768P.
XX 20-MAY-1999; 99US-0134941P.
XX 21-MAY-1999; 99US-0135124P.
XX 21-MAY-1999; 99US-0135353P.
XX 24-MAY-1999; 99US-0135629P.
XX 25-MAY-1999; 99US-0136021P.
XX 27-MAY-1999; 99US-0136392P.
XX 28-MAY-1999; 99US-0136782P.
XX 01-JUN-1999; 99US-0137222P.
XX 03-JUN-1999; 99US-0137528P.
XX 04-JUN-1999; 99US-0137502P.
XX 07-JUN-1999; 99US-0137724P.
XX 08-JUN-1999; 99US-0138094P.
XX 10-JUN-1999; 99US-0138540P.
XX 10-JUN-1999; 99US-0138847P.
XX 14-JUN-1999; 99US-0139119P.
XX 16-JUN-1999; 99US-0139452P.
XX 16-JUN-1999; 99US-0139453P.
XX 17-JUN-1999; 99US-0139492P.
XX 18-JUN-1999; 99US-0139454P.
XX 18-JUN-1999; 99US-0139455P.
XX 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 21-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148664P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
```

PR	20-AUG-1999;	99US-0149722P.	
PR	20-AUG-1999;	99US-0149723P.	
PR	20-AUG-1999;	99US-0149929P.	
PR	23-AUG-1999;	99US-0149902P.	
PR	23-AUG-1999;	99US-0149930P.	
PR	25-AUG-1999;	99US-0150566P.	
PR	26-AUG-1999;	99US-0150884P.	
PR	27-AUG-1999;	99US-0151065P.	
PR	27-AUG-1999;	99US-0151066P.	
PR	27-AUG-1999;	99US-0151080P.	
PR	30-AUG-1999;	99US-0151303P.	
PR	31-AUG-1999;	99US-0151438P.	
PR	01-SEP-1999;	99US-0151930P.	
PR	07-SEP-1999;	99US-0152363P.	
PR	10-SEP-1999;	99US-0153070P.	
PR	13-SEP-1999;	99US-0153758P.	
PR	15-SEP-1999;	99US-0154018P.	
PR	16-SEP-1999;	99US-0154039P.	
PR	20-SEP-1999;	99US-0154779P.	
PR	22-SEP-1999;	99US-0155139P.	
PR	23-SEP-1999;	99US-0155486P.	
PR	24-SEP-1999;	99US-0155659P.	
PR	28-SEP-1999;	99US-0156458P.	
PR	29-SEP-1999;	99US-0156596P.	
PR	04-OCT-1999;	99US-0157117P.	
PR	05-OCT-1999;	99US-0157753P.	
PR	06-OCT-1999;	99US-0157865P.	
PR	07-OCT-1999;	99US-0158029P.	
PR	08-OCT-1999;	99US-0158232P.	
PR	12-OCT-1999;	99US-0158369P.	
PR	13-OCT-1999;	99US-0159281P.	
PR	13-OCT-1999;	99US-0159294P.	
PR	13-OCT-1999;	99US-0159295P.	
PR	14-OCT-1999;	99US-0159329P.	
PR	14-OCT-1999;	99US-0159330P.	
PR	14-OCT-1999;	99US-0159331P.	
PR	14-OCT-1999;	99US-0159637P.	
PR	14-OCT-1999;	99US-0159638P.	
PR	18-OCT-1999;	99US-0159584P.	
PR	21-OCT-1999;	99US-0160741P.	
PR	21-OCT-1999;	99US-0160767P.	
PR	21-OCT-1999;	99US-0160768P.	
PR	21-OCT-1999;	99US-0160770P.	
PR	21-OCT-1999;	99US-0160814P.	
PR	21-OCT-1999;	99US-0160815P.	
PR	22-OCT-1999;	99US-0160980P.	
PR	22-OCT-1999;	99US-0160981P.	
PR	22-OCT-1999;	99US-0160989P.	
PR	25-OCT-1999;	99US-0161404P.	
PR	25-OCT-1999;	99US-0161405P.	
PR	25-OCT-1999;	99US-0161406P.	
PR	26-OCT-1999;	99US-0161359P.	
PR	26-OCT-1999;	99US-0161360P.	
PR	26-OCT-1999;	99US-0161361P.	
PR	28-OCT-1999;	99US-0161920P.	
PR	28-OCT-1999;	99US-0161922P.	
PR	28-OCT-1999;	99US-0161993P.	
PR	29-OCT-1999;	99US-0162142P.	

Query Match 10.9%; Score 116.6; DB 3; Length 1292;  
Best Local Similarity 48.3%; Pred. No. 4.4e-23;  
Matches 434; Conservative 0; Mismatches 444; Indels 21; Gaps 3;

Qy	82	GTCTCTACTTCAACAATGGCGGTAAAGAGATTGGTTCTCTACGTTTCTTGAAGAACTGCAG	141
Db	191	GACITTTACTACGAAAGGCGGTAAAGACATCGCTCGAAACCTTGGTTTCAGCTTTGAG	250
Qy	142	GCTTTCTCTGTTATCTTCATTCTCTCTTACATTACCGCGGCGGAGAGCAACA	201
Db	251	GGTTTCTTTTAAACCCCTTCCCTTGCTATTAATTAAGCCCTGAGCCGTCGAAGACTAAA	310
Qy	202	ATGTGGGTGATGATACAGTTTCTTTCTTATCAAAACCGCGTCTTCTTATCCGCCGTGTTA	261

Db	311	CCATTACCAAAAAAATACTTCTTCTCTTGACA-----CTATCTTTAGTGATATTG	364
Qy	262	TTGTAGGCATCTCTCAGGGTTTGATAACTACTTGTGATGATATGATATAGCTATATCTTC	321
Db	365	GACTTGGCTTGTGTGTCTGGACAATTGATTTTGTACTCATTTGGGCTACTTTACCTTC	424
Qy	322	CAGTTTCTACAGCTGCTCTTATCATTTGCTTTCACAGTTTGTAGCTTTTATAGCTATCTTCAT	381
Db	425	CTGTCTCAACTTCTCTCTTTGATCTCTGCGTCGCAATTTGGCTTTTAAACCGCTCTCTCTT	484
Qy	382	TCCTCATGGTTAAACATAAAGTTCACCTCTTTTACCATTCAATGCTGTGTGTGTGACTG	441
Db	485	ACTTCTTAACACTCACAAAAAATCACACCAITTTATACTCAANTTCACTTGTCTCTTAACCA	544
Qy	442	TTGGTGTCTGGGTTTGGGAAATGCATACCGAAACTGA-----TAAGCCAGTTC	489
Db	545	TATCTTCTACACTTCTTGTGTTATCCAACATGAACAGCAATCTCCCTCTTCTACTTCAAGT	604
Qy	490	ATGAGACTCAACAGCAGTACATAACTGGTTTCTTGATTACTGTAGCAGCAGCTGTTATGT	549
Db	605	CCGAGCCAAAGTCCAAGTATGTGATTGGATACATCTGCGCGGTGGTAGCTAGCTGGTT	664
Qy	550	ATGCTTTTCATCTTGCCATTAGTGGAACTTGCTTACCAGAAAGCTTAAGCAAAACCATGAGCT	609
Db	665	ATTCTCTGGTCTTCTTTAACAGATTACGGGTTTCGAAAGATTCTAAAGAAATACACAT	724
Qy	610	ATACCTTGTGCTCGAGTCCAGTTGATTTTGTGCTCTCCTTGTCTTATTTGTTCAGCGTCA	669
Db	725	TCAAGGCTATTTTAGACATGGCCACATATCCGTCTATGGTAGCTACTTGTGTAGTTGTGG	784
Qy	670	TCGSTATGTTTCATCGCTGGTGATTTCAGAGCAGGCTTTACCAAAAGAACGAAGAGAGTTCA	729
Db	785	TAGGACTTTTGGAAAGTGGTGGTGGGAAAAG---CTGAGTACAGAAATGGAAGAGTTTC	841
Qy	730	AGCTTGGAGAGGCATTGTTCTATGTGTGGCTGTGTGTTTTTCAGCCATCATATGGCAAGGCT	789
Db	842	AACTAGGGAAGCTCATACATTTTGATAAAACATCGGTTTCAACGATATATGGCAAGCTT	901
Qy	790	TCCTTCTGGAGGCCATTTGGATTAAATCTTCTCCACATCGTCTCTCGTCTCGGGTATTAGA	849
Db	902	GTTTGTATTGGAAGTGTGGTTTGATTATCGAAGTTTCATCGCTTTTTCAAATGTCATAA	961
Qy	850	TATCAGTGTCTTTTGCCCAATTACAGAGGTTTTAGCTGTTATATTCTACCATGAAAAGTTTC	909
Db	962	GCACTCTTTGTTTACCAAGTTGCGCTGTCTTGCTGTGTGTTCTTCTCCGTAAGAGATGA	1021
Qy	910	AAGCTGAGAAGGAGACTTTCTCTTGTCTCTCTCCCTTTGGGGCTTTGTCTCTTACTTTTAT	968
Db	1022	GTGGAATCAAGTTGGTTTGCAATGTTTTTGGCCATCTGGGGAATTTGTTTCTTATGTTAT	1080

Search completed: November 1, 2004, 07:20:11  
Job time : 566.296 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 1, 2004, 05:45:29 ; Search time 3851.17 Seconds  
(without alignments)  
10228.408 Million cell updates/sec

Title: US-09-913-767-6  
Perfect score: 1081  
Sequence: 1 caaatccaacagtccaagat.....ataagctgatatttgaaag 1081

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_hic:\*  
4: gb\_est3:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_gss1:\*  
9: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	858.8	79.4	1519	CNS0A3B5	BX826795 Arabidops
2	433.6	40.1	822	BZ491815	BZ491815 BONPU42TF
3	425.4	39.4	926	BX835351	BX835351 BX835351
4	424.2	39.2	849	BX838927	BX838927 BX838927
5	369.6	34.2	1340	CNS0A29J	BX827536 Arabidops
6	369.2	34.2	668	BQ704880	BQ704880 Bn01_01n0
7	366.6	33.9	1375	CNS0A2S8	BX827187 Arabidops
8	363.4	33.6	1431	CNS0A2XR	BX829081 Arabidops
9	357.2	33.0	1312	CNS0A2U8	BX827470 Arabidops
10	331.6	30.7	1359	CNS0A2R3	BX827477 Arabidops
11	299.8	27.7	768	AJ795195	AJ795195 AJ795195
12	284.2	26.3	1323	CNS0A3GP	BX826811 Arabidops
13	279.4	25.8	666	BQ157452	BQ157452 NF105A121
14	269.8	25.0	922	BZ967767	BZ967767 PUDG567D
15	263	24.3	782	CO105310	CO105310 GR_Bb003
16	259	24.0	480	AV522367	AV522367 AV522367
17	256.4	23.7	627	BQ155273	BQ155273 NF078F121
18	255.6	23.6	770	CG344371	CG344371 OG4AG15TC
19	253.4	23.4	1374	CNS0ABEV	BX814955 Arabidops
20	249.6	23.1	837	CB292743	CB292743 UCRCS01_0
21	247.8	22.9	805	BX927581	BX927581 BX927581
22	247.4	22.9	699	B77316	B77316 T3213TF TAM
23	242.4	22.4	691	BZ087060	BZ087060 lkh10d12.
24	240	22.2	399	AV559386	AV559386 AV559386

ALIGNMENTS

RESULT 1	LOCUS	CNS0A3B5	1519 bp	mRNA	linear	HTC	06-FEB-2004
DEFINITION	Arabidopsis thaliana Full-length cDNA Complete sequence from clone GSLTFB632D12 of Flowers and buds of strain col-0 of Arabidopsis thaliana (thale cress).						
ACCESSION	BX826795						
VERSION	BX826795.1	GI:42460970					
KEYWORDS	HTC; GSLT cDNA.						
SOURCE	Arabidopsis thaliana (thale cress)						
ORGANISM	Arabidopsis thaliana						
REFERENCE	Castelli, V., Aury, J.M., Jaillon, O., Wincker, P., Clepet, C., Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V., Temple, G., Caboche, M., Weissenbach, J., and Salanoubat, M.						
AUTHORS	Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences: A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation						
TITLE	Unpublished						
JOURNAL	2 (bases 1 to 1519)						
REFERENCE	Direct Submission						
AUTHORS	Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)						
JOURNAL	The sequences are based on single pass reads.						
COMMENT	Life Technologies (a division of Invitrogen) members carried out full-length libraries construction : Temple G. Genoscope members carried out sequencing and annotation : Castelli V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M. URV INRA : Clepet C., Caboche M. Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich Information center for Protein Sequences). 5 prime and 3 prime are assembled with Phrap. http://www.genoscope.cns.fr/externe/sequences/banque_projet_BF/Full length http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis. Location/Qualifiers						
FEATURES	1..1519						
source	/organism="Arabidopsis thaliana"						
	/mol_type="mRNA"						
	/strain="Col-0"						
	/db_xref="taxon:3702"						

BG525645 52-8 Stev  
CO105307 GR\_Bb003  
BX87065 Arabidops  
AJ795240 AJ795240  
BQ121182 EST606758  
CF069001 EST669722  
BM176948 saJ74h01.  
CK759270 pam01-13m  
CF445584 EST681929  
BX826394 Arabidops  
BM408950 EST583277  
AJ729914 BNLGH1568  
AJ800150 AJ800150  
BI310188 EST531193  
BM413141 EST587468  
CA255028 SCEPFL417  
CK096760 UB25CPFL10  
CNI46235 WOUND1\_39  
CA920055 EST637773  
CNI25253 RHOH1\_10  
AI777352 EST263760

25 233.2 21.6 630 4 BG525645  
26 233 21.6 739 7 CO105307  
27 221.4 20.5 1289 3 CNS0A3HI  
28 220.4 20.4 765 1 AJ795240  
29 218.6 20.2 762 5 BQ121182  
30 217.2 20.1 800 6 CF069001  
31 214.6 19.9 568 4 BM176948  
32 213.4 19.7 562 7 CK759270  
33 206.4 19.1 809 3 CF445584  
34 205.4 19.0 1299 3 CNS0A41J  
35 203.4 18.8 735 4 BM408950  
36 203 18.8 509 1 AJ729914  
37 202.2 18.7 684 1 AJ800150  
38 201.8 18.7 806 4 BI310188  
39 200.2 18.5 762 4 BM413141  
40 193.4 17.9 652 6 CA255028  
41 192.6 17.8 589 7 CK096760  
42 191.2 17.7 765 7 CNI46235  
43 190.2 17.6 676 6 CA920055  
44 189.6 17.5 761 7 CNI25253  
45 189.2 17.5 633 1 AI777352

Db	1030	AATAAAGCGTTTGGAACTCCAGTGGTTCCTATCTTGGCTGTAATCATTTTCCATGACA										1089										
Qy	896	AAATGAATGGGTTAAAGGTGATTTCTATGATCCCTAGCTATTTGGGGTTTTCACTTCTCTATG										955										
Db	1090	AATAAATGGTTAAAGGTGATTTCTATGATCCCTAGCTATTTGGGGTTTAACTTCTCTATG										1149										
Qy	956	TCTACCAACAATATCTTGATGACAAAACCTTGAAGAAAATATCATGAAATCACAACAACAG										1015										
Db	1150	TCTACCAAAAATATCTTGATGACAAAACCTTGAAGAAAATCATGAAATACCAACAACAG										1209										
Qy	1016	AATCCCTGACCCACCAAGACGAGAGTCAACTTGGCAATCAAAATAAGCTGATATTT										1075										
Db	1210	AATCCCTGACCCACCAAGACGAGAGTCAACTTGGCAATCAAAATAAGCTGATATTT										1269										
Qy	1076	TGAAG 1081																				
Db	1270	TGAAG 1275																				
RESULT 2																						
BZ491815/c																						
LOCUS																						
DEFINITION				BZ491815	822 bp	DNA	linear	GSS 16-DEC-2002														
				BONPU421F BO.1.6.2_KB_tot Brassica oleracea genomic clone BONPU42,																		
ACCESSION				genomic survey sequence.																		
VERSION				BZ491815																		
KEYWORDS				BZ491815.1 GI:27000142																		
SOURCE				GSS.																		
ORGANISM				Brassica oleracea																		
REFERENCE				Brassica oleracea																		
AUTHORS				Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.																		
TITLE				1. (bases 1 to 822)																		
JOURNAL				Town, C.D., Van Aken, S., Utterback, T., Koo, H. and Fraser, C.M.																		
COMMENT				Whole genome shotgun sequencing of Brassica oleracea																		
				Unpublished (2001)																		
				Contact: Chris Town																		
TIGR																						
				9712 Medical Center Drive, Rockville, MD 20850, USA.																		
				Tel: 301-838-3523																		
				Fax: 301-838-0208																		



TITLE Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.  
Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences:  
A Combined Approach to Evaluate and Improve Arabidopsis Genome  
Annotation

JOURNAL  
COMMENT

Unpublished (2004)  
Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: [secref@genoscope.cns.fr](mailto:secref@genoscope.cns.fr), Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr)

The sequences are based on single pass reads.

Life Technologies (a division of Invitrogen) members carried out

full-length libraries construction : Temple G.

Genoscope members carried out sequencing and annotation : Castelli

V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,

Schachter V., Weissenbach J., Salanoubat M.

URGV INRA : Clepet C., Caboche M.

Annotation is based on the June 2003 version of the Arabidopsis

genome released by MIPS (Munich Information center for Protein

Sequences).

[http://www.genoscope.cns.fr/externe/sequences/Banque\\_Projet\\_EF/EST](http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/EST)

<http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis>.

## FEATURES

source

1. 849

/organism="Arabidopsis thaliana"

/mol\_type="mRNA"

/ecotype="Col-0"

/db\_xref="taxon:3702"

/clone="GSLTFB632F12"

/tissue type="Flowers and buds"

/clone\_lib="Arabidopsis thaliana Flowers and buds Col-0"

## ORIGIN

Query Match 39.2%; Score 424.2; DB 5; Length 849;  
Best Local Similarity 80.6%; Pred. No. 5.9e-104;  
Matches 623; Conservative 0; Mismatches 78; Indels 72; Gaps 8;

Qy 1 CAATCCACAGTTCAAGATGAAGAAATTCAGTCAGTACGAGCCCAAGCAAGTATCTC 60  
Db |||||||  
Qy 86 CAATCCACAGTTCAAGATGAAGAAATTCAGTCAGTACGAGCCCAAGCAAGTATCTC 145  
Db |||||||  
Qy 61 ACTCTAACACATACAAACGGTGGCTCAG----- 88  
Db |||||||  
Qy 146 ACTCTAACACATACAAACGGTGGCTCAGGGTGACTTTATATACATCTTTGTCTATTTCAG 205  
Db |||||||  
Qy 89 -----GAGTATCTATGACACAGCGAGGAAACAGTAA 120  
Db |||||||  
Qy 206 GCCAAACAGTTGCTACAAATTTGGCGCAGAGTACTATGACACGGAGGAAACAGTAA 265  
Qy 121 TGGCTAGCAACGGTAGTTCAACTGTTGGCTTTCTCTGCTACTTCCATATTATATCTTG 180  
Db |||||||  
Qy 266 TGGCTAGCAACGGTAGTTCAACTGTTGGCTTTCTCTGCTACTTCCATATTATATCTTG 325  
Db |||||||  
Qy 181 TCATTTAAACACATGCAACACTGATAGAGATGGAAGAAACCTCACCTTA--GGAACC 238  
Db |||||||  
Qy 326 TCATTTAAACACATGCAACACTGATAGAGATGGAAGAAACCTCACCTAGTGAATCC 385  
Qy 239 GTGATATGGTTACGTAGTGGTTGAC--TTCTGTAGGAGCAGATGCTATCTGTACTCC 297  
Db |||||||  
Qy 386 GTGATATGGTTACGTAGTGGTTGAC--TTCTGTAGGAGCAGATGCTATCTGTACTCC 445  
Qy 298 ATTGACATCTTTACTTACCGGTTTCTACCTATTCCTGATCTGTCATCTCAGTTAGCC 357  
Db |||||||  
Qy 446 TTGCTGCTCTTACTTTCCCTTTTCTCC---TTTCTGATCTGCTCTCTGTGATCC 502  
Db |||||||  
Qy 358 TTCAATGCTTCTTCTTCTTATTTCTTAACACAAAAACTTACCCCTPATCATTTTAAAT 417  
Db |||||||  
Qy 503 TTAAA--GTGCTACGCTTTTGCGTAACTCAC-AAAACTTACCCCTPATCATTTTAAA- 558  
Db |||||||  
Qy 418 TCTCTTTTCCCTTAACTATATCTTCCACCCTACTTGCAATTAATGAGGAGACAGAC 477  
Db |||||||  
Qy 559 TCTCTTTTCCCTTAACTATTA--CTCTCTCCGACTTGTATTCATTAATGAGGAGACATTC 616  
Db |||||||  
Qy 478 TCCACAAAGTTTACAAAGGAGATGTGTCAAAGGTTTTCATATGCAACCGTTGCTGCGTCT 537  
Db |||||||

Db 617 TCATATAGAGATAAAGGATATTATTTTAAGGTTACAGATGCACCGTGTGCGCCC 676  
Qy 538 GCTGGTTATGCTAGTCTTTATCCCTACACAGCTAGCCTTTCTAAAAGTCTTAAAGAAG 597  
Db |||||||  
Qy 677 GCGGTAATGCTCGAGTCTTATCCGACAATATCCAGCCTTTCTAAAAGTCTTAAAGAAG 736  
Db |||||||  
Qy 598 CAAATTTCTCAGNAGTTATGGATATGATATATCTACGTGAGTCTAGTGCCAGTTGTGTT 657  
Db |||||||  
Qy 737 CATATTTCTCATATGATGGATTTGATAATCTTTGTGATACACAGTGGCTGTAGAGTT 796  
Db |||||||  
Qy 658 AGCGTGTGGGGCTTTTGTAGCAGTACGTGGAACACTTTGAGCAGCTGAAAT 710  
Db |||||||  
Qy 797 AGCGTGGAGGGGCTTTAGTACGACGAGTGGGATACTTTGAGTAGTGTAAAT 849  
Db |||||||

## RESULT 5

CNS0A29J

LOCUS

DEFINITION

Arabidopsis thaliana Full-length cDNA Complete sequence from clone

GSLTUS712A05 of Adult vegetative tissue of strain col-0 of

Arabidopsis thaliana (thale cress).

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

The sequences are based on single pass reads.

Life Technologies (a division of Invitrogen) members carried out

full-length libraries construction : Temple G.

Genoscope members carried out sequencing and annotation : Castelli

V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,

Schachter V., Weissenbach J., Salanoubat M.

URGV INRA : Clepet C., Caboche M.

Annotation is based on the June 2003 version of the Arabidopsis

genome released by MIPS (Munich Information center for Protein

Sequences). 5 prime and 3 prime are assembled with Phrap.

[http://www.genoscope.cns.fr/externe/sequences/Banque\\_Projet\\_EF/Full](http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full)

length

<http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=arabidopsis>.

Location/Qualifiers

1. .1340

/organism="Arabidopsis thaliana"

/mol\_type="mRNA"

/strain="Col-0"

/db\_xref="taxon:3702"

/clone="GSLTUS712A05"

/tissue type="Adult vegetative tissue"

/plasmid="pCMVSPORT\_6"

1. .1340

/gene="At4g18197"

ORIGIN

Query Match 34.2%; Score 369.6; DB 3; Length 1340;

Best Local Similarity 62.2%; Pred. No. 4.6e-89;

Matches 601; Conservative 0; Mismatches 359; Indels 6; Gaps 1;

QY	89	GAGTATAC	TATGAC	AAACGG	AGAAACAG	TAAATGG	CTAG	CAACCG	TAGTTC	CAACTG	TGTG	148
DB	225	GATTGTAC	TATG	AAATAG	GAAGAAATAG	CACATATG	TG	GTAA	CACTCT	TTCA	CACTG	284
QY	149	GCTTTC	TG	TGCTACT	CCATAT	TATATCT	TG	TCATTTA	-----	AAACAC	ATCA	202
DB	285	GCTTCC	CTG	TACTG	TTCTG	TTCG	CGCTTCT	TTTCG	GAATC	AGGCA	CCAAATCA	344
QY	203	CTGAT	PAG	AGATG	GGAAAA	AGAACCT	CACCT	AG	GAACCG	TGTAT	TGGTTT	262
DB	345	ATACAA	ATG	T	CAGTCAG	TCCCTT	CCAC	CCCTT	GCATCG	GT	TACTTTG	404
QY	263	GACTT	CT	TG	TAGG	ACAGAT	TGCTAT	CTG	TACT	CCAT	TGGACTT	322
DB	405	GACTG	CT	AGT	TCCGCTT	ATGCTT	ATTTG	CT	GCAG	TAGG	TGCTCT	464
QY	323	CTAC	CTAT	TCCCT	GTATCT	GTGCAT	CTCAG	TTAG	CCCTCA	ATGCTT	CTTCTCT	382
DB	465	CTACT	TT	CTCCCT	CATCTT	GGCCCTC	ACAG	TTGG	CCCTT	CAC	TGCTTTT	524
QY	383	TTAACT	CA	CAAAA	ACTT	ACCCCT	ATCA	TTTTAA	ATCTCT	TTTCT	CTTAACT	442
DB	525	TTAACT	CG	CAAA	AGTTCA	CTCTT	TG	ATG	TAGTCAG	TTCTT	TGCTTCT	584
QY	443	CCAC	CT	TACT	GCAT	TTCA	ATATG	AGG	AGACAG	ACTCC	AAAAAG	502
DB	585	CTGCT	CT	TTCT	TG	TGGTCA	CACTG	ATT	CAG	AAAACT	CAATAT	644
QY	503	ATGTC	AAAG	TTT	CATATG	CA	CCGTTG	CTG	CGCTG	CTG	TGGTTAT	562
DB	645	ATG	TG	ATCGGG	TT	CATATG	TACCA	TCG	TGCTT	CCGCT	TGGAT	704
QY	563	TACA	AC	AGCT	AG	CTTTCT	TA	AAAGT	CTTAA	AGAG	CAAAAT	622
DB	705	TGAT	AC	AAATG	CTTT	CAGG	AAAG	TTTT	CAC	AGAC	ATA	764
QY	623	TGAT	AA	ATCT	AC	GT	GAGT	CT	AGT	CGCC	AGTTG	682
DB	765	TGGC	AT	TTAC	CA	GTCT	CTAG	TTG	CGAG	TTCT	CATAG	824
QY	683	GTG	AGT	GGAAAA	CTTT	GAG	CA	GTG	AAAT	TG	ATACT	742
DB	825	GAG	AGT	GGAAAA	CTTTG	CA	AGT	GAG	ATG	AGAAA	CTAC	884
QY	743	TTAT	GA	ACCT	AGT	GTG	GAC	AGT	GT	TACT	CG	802
DB	885	TTTT	GA	CTTT	AG	CC	TG	CGG	CA	GTCTA	CCT	944
QY	803	TGAT	CT	TCG	AG	CTCT	CTCT	AT	TTCT	CAAA	TGA	862
DB	945	TGAT	CT	TCG	AGT	CACT	CTCT	GTG	TTCT	CCA	ATT	1004
QY	863	TTCC	TA	CT	TG	GGT	TA	ATCA	TTTT	CA	ATG	922
DB	1005	TTCC	AG	TTG	CGCG	AGT	GATAG	TTTT	TCC	ATG	ATAG	1064
QY	923	TGAT	CC	TAG	CT	AT	TTTGGG	TTT	CAC	TTTCT	ATG	982
DB	1065	TTAT	TTT	TAG	CT	AT	CTG	GGCTT	CTCT	CA	TTTCT	1124
QY	983	ACTT	GA	AGAAAA	AATCAT	GAAAT	CA	CAAC	AG	AAAT	CCCT	1042
DB	1125	AGTT	GA	TACT	AG	CCAC	CA	CA	AGT	TGCT	GTAG	1184
QY	1043	AGT	CAA	1048								
DB	1185	AC	CAA	1190								
RESULT 6												
BQ704880												
LOCUS												
BQ704880 668 bp mRNA linear EST 16-JUL-2002												

RESULT 6  
BQ704880  
LOCUS

DEFINITION Bn01\_01n06\_A  
Bn01\_AAPC\_ECORC transgenic\_Brassica napus\_overexpressing\_BNCBF17\_co  
natitatively\_frost\_tolerant Brassica napus cDNA clone Bn01\_01n06,  
mRNA sequence.

ACCESSION B0704880  
VERSION B0704880.1 GI:21844299  
KEYWORDS EST.  
SOURCE Brassica napus (rape)  
ORGANISM Brassica napus  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.  
1 (bases 1 to 668)  
REFERENCE Singh,J., Allard,G., Tinker,N., Robert,L., Lacroix,C., De Moors,A.,  
Chagnon,J., Farah,S., Couroux,P. and Hattori,J.  
AUTHORS Expressed Sequence Tags from constitutively frost tolerant  
TITLE transgenic Brassica napus overexpressing BNCBF17  
JOURNAL Unpublished (2002)  
COMMENT Contact: Singh,J.A.  
Eastern Cereal and Oilseed Research Centre  
Agriculture and Agri-food Canada  
KW Neatby Bldg., Central Experimental Farm, Ottawa, Ontario, K1A  
0C6, Canada  
Tel: (613) 759-1662  
Fax: (613) 759-1701  
Email: singhja@agr.gc.ca.  
Location/Qualifiers  
1. .668  
/organism="Brassica napus"  
/mol\_type="mRNA"  
/cultivar="Westar"  
/db\_xref="taxon:3708"  
/clone="Bn01\_01n06"  
/tissue\_type="fourth leaf"  
/dev\_stage="3 weeks seedling grown at room temperature"  
/clone\_lib="Bn01\_AAPC\_ECORC transgenic\_Brassica napus over  
expressing\_BNCBF17\_constitutively\_frost\_tolerant"  
/notes="vector: Bluescript SK-/XhoI-EcoRI; Site 1: EcoRI;  
Site 2: XhoI; Germinated in soil flats and seedlings grown  
for 3 weeks in a Conviron E-15 cabinet set at 20°C/16 hr  
light (250 Em-2sec-1) and 16 °C/ 8 hr dark. Fourth leaves  
collected at 9 am and immediately frozen."

ORIGIN

Query Match 34.2%; Score 369.2; DB 5; Length 668;  
Best Local Similarity 79.0%; Pred. No. 5e-89;  
Matches 433; Conservative 22; Mismatches 90; Indels 3; Gaps 2;

Qy 524 CCGTTGCTGGCTGCTGGTTATGGTCTAGTCTTATCCCTACACAGTAGCTTCTTAA 583  
Db 1 CCATTGGTGCTTCTGCTGGTATGGTCTAAGCTTATCTTYACAAAGTAGTCCAGTA 60

Qy 584 AAGTCTTAAGAAGCAAAATTTCTCAGAAGTTATGGATATGATAATCTACGTGAGTCTAG 643  
Db 61 AAGTCTCTAAAGAGGCAAACTTCTCASAAGTCATGGAATTSATCATCTACGTAAAGTCTAG 120

Qy 644 TGSCCAGTTGTGTTAGCTGGTGGGGCTTTTGTCTAGCAGTAGTGCAAACTTTTCAGCA 703  
Db 121 TGSCCMGCTGTGTTAAGCGTGGTGGGGCTCTTTGCTAGTGGAGTGGGAACTTTSMGCA 180

Qy 704 GTGAATGGATACACTACAAACATGGGAAGGTATCTCTACATTATGAACCTAGTGTGGACAG 763  
Db 181 GTCAATGGATAGCTACAAAGCTTGGAAAGGTATCTTACGTTATGAACCTAGTGTGGACGG 240

Qy 764 CTGTTACCTGGCAGTTATTCTCCATCGTGGGCACAGGACTGATCTTTCGAGCTCTCT -CT 822  
Db 241 CTGTTACCTGGCAGGTATTCAACATTTGGTAGCACGTGGCTTATCTTCGAGATTTCTCTCT 300

Qy 823 CTATTCTCAAAATGCATAG -CGTTTGGGACTCCAGTGGTTCCTATCTTGGCTGTAA 880  
Db 301 CTATTCTCAAAAGCTATGCGGTGTTGGGTTTGGCTCTGGCGYCCGCACTSGCTGTCA 360

Qy 881 TCATTTTCCATGACAAAATGAATGGGTAAAGGTGATTTCTATGATCTTAGCTATTTTGGG 940





VERSION	BX829081.1	GI:42460228
KEYWORDS	HTC; GSLT_cDNA.	
SOURCE	Arabidopsis thaliana (thale cress)	
ORGANISM	Arabidopsis thaliana	
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.	
AUTHORS	Castelli,V., Aury,J.M., Jaillon,O., Wincker,P., Clepet,C., Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V., Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.	
TITLE	Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences: A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation	
JOURNAL	Unpublished	
REFERENCE	2 (bases 1 to 1431)	
AUTHORS	Genoscope.	
TITLE	Direct Submission	
JOURNAL	Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)	
COMMENT	The sequences are based on single pass reads. Life Technologies (a division of Invitrogen) members carried out full-length libraries construction : Temple G. Genoscope members carried out sequencing and annotation : Castelli V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M. URGV INRA : Clepet C., Caboche M. Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich Information center for Protein Sequences). 5 prime and 3 prime are assembled with Phrap. http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_BF/Full length http://www.genoscope.cns.fr/cgi-bin/ggb?ggb?source=Arabidopsis.	
FEATURES	Location/Qualifiers	
source	1..1431	
	/organism="Arabidopsis thaliana"	
	/mol_type="mRNA"	
	/strain="Col-0"	
	/db_xref="taxon:3702"	
	/clone="GSLTSL692803"	
	/tissue_type="Silique"	
	/plasmid="pCMVSPORT_6"	
Gene	1..1431	
	/gene="At4g18205"	
ORIGIN		
Query Match	33.6%;	Score 363.4; DB 3; Length 1431;
Best Local Similarity	62.5%;	Pred. No. 2.3e-87;
Matches	587; Conservative	0; Mismatches 346; Indels 6; Gaps 1;
QY	59	TCACTCTAACACATACAAAGGGTGGCTCAGAGTATATGATGACGAGGAAACAGTA 118
Db	279	TCGCGAACCACTTGCTACAGTTCTGGGTAGACTGTACTATGAAATGAGGAAAGCA 338
QY	119	AATGGCTAGCAACGGTAGTTCAACTTGTGGCTTCCCTGTGCTACTTCCATATATCT 178
Db	339	CATATGTGTAACACTTCTCACTCATTTGGCTTCCCTGTGCTACTGATCTCTCCGGTCT 398
QY	179	TGTCATTTA-----AAACACATGCAACACTGATAGATGGAAGAAAGCACTCACCTA 232
Db	399	TTTCTCGAATCAGGCAACCAATCAACAGATACAAATTCAGTCAGTCCCTTCCTTCA 458
QY	233	GGAACCGTGTATGGTTAGCTAGTCTGGACTTCTTCTAGGACGAGATTGCTATCTGT 292
Db	459	CCACCTTGATCGGTTTACTTGTGCAGTGGACTGTGCTGTCTTATGCTTATTTGT 518
QY	293	ACTCCATTGGACTCTTTTACTTACCGTTTCTACCTATTCCCTGATCTGTGCTATCTCAGT 352
Db	519	CTGCAGTTGGTGTCTTACTTACCACTCTCTACTTCTCCCTCATCTTGGCTCACAGT 578
QY	353	TAGCCTTCAATGTTTCTCTTATTTCTTAATCTCAACAAACCTTACCCCTATCATTT 412
Db		

Db	579	TGGCCTTCACTGCTCTTTTCTCATATTTCTTAACTCGCAAAAGTTTCACTCTCTTTGATAG 638
QY	413	TAAATTCCTTTTCTCTTAACTATATCTTCAACCTCTCTGCTTCAATTAATAGGAGAGA 472
Db	639	TCATTTCTTTGTTTCTCTTACGGTTTCTCTGCTCTCTCTGCTGCTCAACACTGATTTCAG 698
QY	473	CAGACTCCACAAAGTTACAAAGGAGAGATATGTCAAAGTTTTCATATATGACCGTTGCTG 532
Db	699	AAACACACAAATATGATCTAGAGTACAATATGTGTTGGTTTCATCTGTACCATTGGTG 758
QY	533	CGTCTGCTGTTATGCTGTAGTCTTATCCCTCAACAGCTAGCTTCTTAAAGTCTCTAA 592
Db	759	CTTCGCTGGGATTGGATTGGTACTATCTCTGATACAACCTGCTCTTCAGGAAAGTTTCA 818
QY	593	AGAACGCAAAATTTCTCAGAAAGTTATGATATATCTACGTGAGTCTAGTGGCCAGTT 652
Db	819	CGAAGCATACCTCTCAGCAGTCTGGACTTGGCCAAATTACCAGTCTCTAGTTGCTACTT 878
QY	653	GTGTTAGCGTGGTGGGCTTTTGTGTCAGAGTGAAGTGAAGAACTTTTGACGAGTGAATGG 712
Db	879	GTGTGCTACTCATAGGACTGTTTGCAGGTGGAGAGTGGAGAAATTCCTGCCAAGTGAGATGA 938
QY	713	ATAACTACAAACATGGGAAGGTATCTCATATTATGAACCTAGTGTGGAGCAGCTGTACCT 772
Db	939	GAACTTACATCTGCGGAAAGTGTATATATCTTGACTTTGGCCTCAGCAGCTATTTTCT 998
QY	773	GGCAGTTATTCTCCATCGGTGCGCAGGACTCATCTTCGAGCTCTCTCTCTATTCTCAA 832
Db	999	GGCAAGTATACACTGTTGGTGTGGGATTAATCTTCAGTCTCTCTGTGTTCTCCA 1058
QY	833	ATGCAATTAAGCGTTTGGGACTCCAGTGGTTCCTATCTTGGCTGTAACTATTTTCCATG 892
Db	1059	ATTCCATAACTCTCTGGGACTGCTATAGTTCCAGTTCTAGCAGTGATAGTTTCCATG 1118
QY	893	ACAAATGAATGGTTAAAGTGATTTCTATGATCTCTAGTATTTGGGGTTTCACTTCTCT 952
Db	1119	ATAAGATGGATGCATCAAAAGATTTTCTCCATCATTTTAGTCTATCTGGGGCTTCTTTCA 1178
QY	953	ATGTCCTACCAACAATATCTTGATGACAAAACTTCAAGA 991
Db	1179	TTGTCATCAGAACTACCTCGACGAAAGAGTTGAAGA 1217

CNS0A2U8	1312 bp	mRNA	linear	HTC 06-FEB-2004
Arabidopsis thaliana Full-length cDNA Complete sequence from clone GSUTLS92F07 of Adult vegetative tissue of strain col-0 of Arabidopsis thaliana (thale cress).				
Accession	BX827470			
Version	1	GI:42460504		
Keywords	HTC; GSLT_cDNA.			
Source	Arabidopsis thaliana (thale cress)			
Organism	Arabidopsis thaliana			
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.			
Reference	1 (bases 1 to 1312)			
Authors	Castelli,V., Aury,J.M., Jaillon,O., Wincker,P., Clepet,C., Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V., Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.			
Title	Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences: A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation			
Journal	Unpublished			
Reference	2 (bases 1 to 1312)			
Authors	Genoscope.			
Title	Direct Submission			
Journal	Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)			
Comment	The sequences are based on single pass reads. Life Technologies (a division of Invitrogen) members carried out			



Query Match		30.7%;	Score 331.6;	DB 3;	Length 1359;
Best Local Similarity		61.1%;	Pred. No. 9.9e-79;		
Matches 591;		Conservative 0;	Mismatches 369;	Indels 8;	Gaps 3;
QY	89	GAGTATACATGACACGAGGAGAAACAGTAATAGGCTAGCAACGGTAGTTCAACTGTTG	148		
DB	199	GATTGACTATGAATGAGGAGAAATAGCACATATGTGTGAACACTTCTTCAACTCATTTG	258		
QY	149	GCTTTCCTGCTGCTACTTCCATATATATATCTTGTCAITTTA-----AAACACATGCAACAA	202		
DB	259	GCTTCCCTGCTACTGTTCTGCTCCGCTTCTTTTCTCGAATCAGGCAACCAATCAACAG	318		
QY	203	CTGATAGAGATGAAAGAAAGAACCTCACCTAGGACCGTGTATTGGTTTACGTAGTCTTG	262		
DB	319	ATACAAATTCAGTCAGTCCCTTCTCTTACCACCCCTTGCAATCGGTTTACTTGTGCACTG	378		
QY	263	GACTTCTTGTAGAGACAGATTGTATCTGTACTCCATTGGACTTCTTTACTTTACCCGTTT	322		
DB	379	GACTGTAGTGTCCGCTTATGCTTATTTGTCTGCAGTAGGGTTGCTCTACTACCAAGTCT	438		
QY	323	CTACCTATTCCCTGATCTGTGATCTCAGTTAGCCTTCAATGCTTTCTTCTCTTATTTTC	382		
DB	439	CTACTTCTCCCTCATCTTGGCCTCAGTTGGCCTTCACTGCTTTTCTCATATTTCC	498		
QY	383	TTAACTCACAAAACTTACCCCTTATCAITTTTAAATTTCTCTTTTCTCTTAACTATATCTT	442		
DB	499	TTAACTCGCAAAAGTTTCACTCCCTTGTATAGTAGTGGTTCTTTGCTTCTCCTCACTGTATCCT	558		
QY	443	CCACCTACTTGCATTCATATATAGGAGACAGACTCCACAAAGTTTACAAAGGAGGT	502		
DB	559	CTGCTCTTCTTGTGGTCAACACTGATTTCAGAAACTCAACTAATGTATCTAGAGTACAGT	618		
QY	503	ATGTCAAAAGTTTTCATATGACCGTTGCTGCGTCTGCTGTTATGCTAGTCTTATCCC	562		
DB	619	ATGTGATCGGTTTCAATATGATACCATCGTGCTTCCGCTGGATTTGAGCTGTACTATCTC	678		
QY	563	TACACAGCTAGTCTTTTCTAAAGTCTTAAAGAGCAAAATTTCTCAGAAAGTTTATGGATA	622		
DB	679	TGAGACAAATGATCTTTCAGGAAAGTTTTCAGGAAGCATATACCTCTCAGCAGTAACGGACT	738		
QY	623	TGATATCTACGTAGTCTAGTGGCCAGTTGTGTAGCGTGTGGGCTTTTGTCTAGCA	682		
DB	739	TGGCCATTTACCACTCTCTAGTGTGCGCGTTGTGTAGTTCTCATAGGACTTTTTTGCAGTT	798		
QY	683	GTGAGTGGAAACCTTTGAGCAGTGAATGATAAATCTACAAACATGCGAAGGTATCCTACA	742		
DB	799	GAGAGTGGAAACCTTTGCGAGTGAAGATAAATCTACAACTCGGGAAGTGTATATG	858		
QY	743	TTATGAACCTAGTGTGGACAGCTGTTTACCTGGCAGTTTATCTCCATCGGTGGCACAGGA-	801		
DB	859	TTTTGACTTAAAGCTCGGCAACTATTTTCATGCAAGTCTACACTCTTGGTCTTGTGGAT	918		
QY	802	CTGATCTTCG-AGCTCTCTCTATTTCTCAATGCAATGAAGTTTGGGACTCCCGAGT	860		
DB	919	ATGATCTTCGCAGTCTATCTCTGTGTTCACCCATTCATCCCGCTGTGGGATGCTTAT	978		
QY	861	GGTTCTCTCTTGGCTGTAATCATTTTCCATGACAAATGAATGGTTTAAAGGTGATTTTC	920		
DB	979	CGTTCCTGCTGGCCGCTGATCGTTTTCATGATGAATGGACGCAATCAAAAATCTTCTC	1038		
QY	921	TATGATCTTACTTATTTGGGGTTTTCATTTCTTCTATGTCTACCAACAATATCTTGTATGACAA	980		
DB	1039	CATCAITTTTATGCTATCTGGGCTTCTCTCAITTCGTCTATCAGAACTACCTCGACGAAA	1098		
QY	981	AAACTTGAAGAAAATATGAATAATCAACAACAGAAATCCCTGACCCACAGAGAGCAGA	1040		
DB	1099	GAACTTGAATACTAGCCACACAAAGTCTGTAGAGATCTTCTATCTACCTGTGTGAGGAGG	1158		
QY	1041	AGAGTCAA	1048		
DB	1159	TCACAAA	1166		

RESULT 11	AJ795195	768 bp	mRNA	linear	EST 11-AUG-2004
LOCUS	AJ795195	Antirrhinum majus whole plant	Antirrhinum majus	cdna clone	
DEFINITION	AJ795195	018_3_07_907, mRNA sequence.			
ACCESSION	AJ795195				
VERSION	AJ795195.1	GI:51110523			
KEYWORDS	EST.				
SOURCE	Antirrhinum majus (snapdragon)				
ORGANISM	Antirrhinum majus				
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; Lamiales; Plantaginaceae; Antirrhineae; Antirrhinum.					
REFERENCE	1 (bases 1 to 768)				
AUTHORS	Zachgo,S., Stuber,K., Saedler,H., Sommer,H. and Schwarz-Sommer,Z.				
TITLE	Antirrhinum EST collection				
JOURNAL	Unpublished (2003)				
COMMENT	Contact: Schwarz-Sommer Z				
	Molekulare Pflanzen-genetik				
	MPI fuer Zuechtungs-forschung				
	Carl-von-Linne Weg 10, D-50829, Germany.				
FEATURES	Location/Qualifiers				
source	1..768				
	/organism="Antirrhinum majus"				
	/mol_type="mRNA"				
	/db_xref="taxon:4151"				
	/clone="018_3_07_907"				
	/tissue_type="whole plant"				
	/clone_lib="Antirrhinum majus whole plant"				
ORIGIN					
Query Match	27.7%;	Score 299.8;	DB 1;	Length 768;	
Best Local Similarity	61.9%;	Pred. No. 3.8e-70;			
Matches 475;	Conservative 0;	Mismatches 292;	Indels 0;	Gaps 0;	
QY	109	GGAAACAGTAATAGCTAGCAACGGTAGTTCAACTTGTGTGGCTTTCCCTGTGCTACTTCCA	168		
DB	2	GGGACACGCAATGGCTAGCCACATTAGTACAGTTATAGACTTTTCCAAATTTCTCGTCCA	61		
QY	169	TATTATATCTTGTCTATTTAAACACATGCAACAACTGATAGAGATGGAAAGAACCTCA	228		
DB	62	TTCCAAATGGTTTCAACACCAACGAGGGCCCTAATGAACCAACAAACAAAGAAATTC	121		
QY	229	CCTAGGACCGTGATTTGGTTTACGTAGTCTTGGACTTCTTGTAGGACAGATTGCTAT	288		
DB	122	CCCACTGGTCTTGTATTAATATATGTTAGCTTCGGCAATTTTGTGCGAGGAATTTCCATG	181		
QY	289	CTGTACTCCATTGGACTTCTTTACTTACCGGTTTCTACCTATTTCCTGTGATCTGTGCACT	348		
DB	182	TTATACAAATGGACTCCAAATACCTTCGGTGACTACGTACACTTTGATCTGCGCTAGT	241		
QY	349	CAGTTAGCCTTCAATGCTTTTCTCTTATTTTCTTAACTCACAAAAACCTTACCCCTATC	408		
DB	242	CGGTTAGGATTCACGCGCTTTTCTCGTTTTCCTCAACAAACAAAGTTCACTCCTTAC	301		
QY	409	ATTTTAAATCTCTTTTCTCTTAACTATATCTTCCACCGCTACTTGCATTCAATATGAG	468		
DB	302	ATTATAAACTCTGTAGTACTCTCTACGATTAGTTTCAATATTACTAGTGTTTCCAAACTGAT	361		
QY	469	GAGACAGACTCCACAAAGTTTACAAAAGGAGATGTCAAAGGTTTTCATATGCACCGTT	528		
DB	362	CCTGGAGACTCCACAAAGCTCGAAAAAAGATATGTCATTTGGCTTTTATCCACTGTC	421		
QY	529	GCTCGCTCTGCTGGTTATGGTCTAGTCTTATCCCTACAAACAGCTAGCCTTTCTTAAAGTC	588		
DB	422	GCGCCTCTGCTGGCTACGGTTAGTCTTTTCAATCACACAGCTCGCCTTCGGTAAGGTC	481		
QY	589	CTAAGAGCAAAATTTCTCAGAGTTATGGATATGATTAATCTACGTGAGCTAGTGGCC	648		
DB	482	ATCAAGATGAACCAATTCGCGGATTCCTGGATTGCAATCTATCAGAAATTTGGTGGCG	541		



```
VERSION B0157452.1 GI:20294511
KEYWORDS EST.
SOURCE Medicago truncatula (barrel medic)
ORGANISM Medicago truncatula

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eusteroids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.

1 (bases 1 to 666)
Torres-Jerez, I., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J.,
Flores, H.R., Inman, J.T., Weller, J.W. and May, G.D.
Expressed Sequence Tags from the Samuel Roberts Noble Foundation
Medicago truncatula irradiated library
Unpublished (2001)
Contact: May GD
Plant Biology Division
The Samuel Roberts Noble Foundation
2510 Sam Noble Parkway, Ardmore, OK 73402, USA
Tel: 580 224 6650
Fax: 580 224 6692
Email: gdmay@noble.org
Insert Length: 666 Std Error: 0.00
Plate: 105 row: A column: 12
Seq primer: TCACACGAGAACAGCTATGAC.

FEATURES
    source
        1..666
            /organism="Medicago truncatula"
            /mol_type="mRNA"
            /db_xref="taxon:3880"
            /clone="NF105A12IR"
            /tissue_type="seedlings"
            /dev_stage="seedling"
            /clone_lib="Irradiated"
            /note="Vector: Lambda Zap; Seedlings were exposed either
            to 100 Gy gamma or 0.5, 1, 5, or 10 kJ/m2 UV irradiation.
            Gamma-irradiated samples were harvested at 6, 12, 24 and
            48 hours after treatment. UV-irradiated samples were
            harvested 24 hours post-treatment. cDNA was prepared from
            polyA+ enriched, pooled samples of equivalent amounts of
            total RNA from each sample. The cDNA was directionally
            ligated into the Uni-Zap XR vector (Stratagene) and
            packaged using the Gigapack III Gold packaging extracts.
            Phagemids containing cDNA inserts were in vivo excised
            from the recombinant Uni-Zap XR vector using ExAssist
            helper phage and the E. coli strain XL1-Blue MRF'
            (Stratagene). Excised plasmids were plated using SOLR
            cells."

ORIGIN
    Query Match 25.8%; Score 279.4; DB 5; Length 666;
    Best Local Similarity 64.7%; Pred. No. 1.3e-64;
    Matches 431; Conservative 0; Mismatches 232; Indels 3; Gaps 1;

QY 270 TGTAGAGCAGATGCTACTGTCTACTCCATGTCAGCTCTTTACTTACCGGTTTCTACCTA 329
DB 1 TTTAGCAGGGGCAATTTGATGTATTCCAAATGAGCTTCTATATCTTTCACTTTCCACCTA 60
QY 330 TTCCCTGATCTGTCATCTCAGTTCAGTTCCTCAATGCTTCTTCTCTTATTTCTTAACCTC 389
DB 61 TTCTTTTACTATGTGCAACACAATTAGCCCTTCAATGAGGATTTTCTTATTTCTTAAATTC 120
QY 390 ACAAAAACCTTACCCCTATCATTTTAAATTTCTCTTTTCCCTCTTAACATATCTTCCACCCCT 449
DB 121 CCAAAAGTTTCACAGCATTTCATATTCATTTCTGTAGTTCTGCTTACTATTTCTGCTGCCCT 180
QY 450 ACTTGCAATTCATATGAGAGAC----AGACTCCACAAAAGTTACAAAAGGAGAGATGTT 506
DB 181 ACTTGCCATCAATPAACGAGATTTCTGAGGACTCGACGAATCTTCTCAAGATAAGCGAGT 240
QY 507 CAAAGGTTTCATATGACACCGTTGCTGCGTCTGCTGTTATGCTAGTCTTATCCCTACA 566
DB 241 ACTTGGATTTCGCTTCCACCCCTGGTGCATCTGTACATTTTCTTTATATATCTTTCCCTGTT 300

QY 567 ACAGCTAGCCTTTCTAAAGTCTCTAAAGAACCAAAATTTCTCAGAAGTTATGGATATGAT 626
DB 301 GCAGCTTTCTTTTGACAAAGTTATAAAAGCGAAACCTTTTCTGTTATATTAGACATGCA 360
QY 627 ATCTACGTGAGTCTAGTGGCCAGTTGTTAGCTGCTGGTGGGCTTTTCTTCTAGCAGTGA 686
DB 361 ATTTTACCCCTCTTTTCGTCGCTACTTGTGCTTGTGTGTTGGTGTGTTTGCAGGTGGAGA 420
QY 687 GTGGAAAACCTTTGAGCAGTGAATTAAGTAACCTTACAAAACATGGAAGGTATCCTACATTAT 746
DB 421 ATGGAAGAGTTTGAATAGGAGATGAAGGAATATGAAGTGGAAAGGTCTCATATATAT 480
QY 747 GAACCTAGTGTGAGCAGCTGTTTACCTGCGAGTTATTTCTCCATCGGTGGCAGGAGTGA 806
DB 481 GACTTTGTTTGGACTGCTGTGACATGCAAGTATCTTCCATTGGTATGTTAGGGTTAGT 540
QY 807 CTTGAGCTCTCTCTCTATTTCTCAATGCAATAGCGTTTGGGAGCTCCAGTGGTTCC 866
DB 541 TTTAGATGATCTTTCATTGTTCTCAATGTATATAGGTACAGTGGTACTTCCCGTAGTCC 600
QY 867 TATCTTGGCTGTAATCATTTTCCATGACAAAATGAATGGTTAAAGGTGATTTCTATGAT 926
DB 601 TATCTTGTCTGTATTTTCTTCCATGACAAAATTAATGGGTGAAGTTTATAGCTTTGAT 660
QY 927 CCTAGC 932
DB 661 GTTAGC 666

RESULT 14
BZ967767
LOCUS BZ967767
DEFINITION BZ967767 922 bp DNA linear GSS 25-MAR-2003
PUDGUS6TD ZM 0.6 1.0 KB Zea mays genomic clone ZMMBTa184115,
genomic survey sequence.
ACCESSION BZ967767.1 GI:29185853
VERSION BZ967767.1
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
Remnick, A., Fraser, C.M., Yuan, Y., San Miguel, P., Ma, J. and
Bennetzen, J.
Maize Genomics Consortium
Unpublished (2003)
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF
Class: sheared ends.
FEATURES
    source
        1..922
            /organism="Zea mays"
            /mol_type="genomic DNA"
            /strain="B73"
            /db_xref="taxon:4577"
            /clone="ZMMBTa184115"
            /clone_lib="ZM 0.6 1.0 KB"
            /note="Vector: PCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
            Cot selected genomic DNA library"

ORIGIN
    Query Match 25.0%; Score 269.8; DB 8; Length 922;
    Best Local Similarity 60.4%; Pred. No. 5.7e-62;
    Matches 445; Conservative 0; Mismatches 292; Indels 0; Gaps 0;

QY 260 TTGGACTTCTTGTAGGAGCAGATTGCTATCTGTACTCCATTGGACTTCTTTACTTACCGG 319
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 1, 2004, 06:20:04 ; Search time 618.418 Seconds  
(without alignments)  
9900.026 Million cell updates/sec

Title: US-09-913-767-5

Perfect score: 1194

Sequence: 1 tcatgagatataataacat.....tgtgatcaagcatatttcc 1194

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3413475 seqs, 253800928 residues

Total number of hits satisfying chosen parameters: 6826950

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*
- 16: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq:\*
- 17: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq:\*
- 18: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*
- 19: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq:\*
- 20: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*
- 21: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1015.8	85.1	3387	9	US-09-938-842A-2212
2	1015.8	85.1	3387	11	US-09-938-842A-2212
3	459.8	38.5	592	9	US-09-770-152-44
4	349.8	29.3	1294	16	US-10-424-599-2524
5	303.8	25.4	2214	16	US-10-424-599-102519
6	279	23.4	1513	17	US-10-437-963-82061
7	247.2	20.7	997	16	US-10-424-599-75161
8	244.2	20.5	2381	17	US-10-437-963-43860
9	215.4	18.0	83698	17	US-10-416-898-9
10	157.2	13.2	522	16	US-10-424-599-114931
11	154.4	12.9	792	16	US-10-424-599-9213
12	146.6	12.3	1386	16	US-10-425-114-12485

13	145	12.1	1779	16	US-10-425-114-28473	Sequence 28473, A
14	141.8	11.9	728	17	US-10-767-701-9000	Sequence 9000, Ap
15	138.8	11.6	747	17	US-10-767-701-6806	Sequence 6806, Ap
16	134.4	11.3	2605	17	US-10-437-963-83900	Sequence 83900, A
17	122.2	10.2	1146	16	US-10-425-114-10516	Sequence 10516, A
18	122.2	10.2	1233	16	US-10-424-599-30274	Sequence 30274, A
19	120	10.1	386	9	US-09-770-791-85	Sequence 85, Appl
20	118.4	9.9	332	11	US-09-732-627A-1196	Sequence 1196, Ap
21	107	9.0	1432	17	US-10-437-963-1897	Sequence 1897, Ap
22	100	8.4	671	17	US-10-767-701-5986	Sequence 5986, Ap
23	93.4	7.8	1333	16	US-10-425-114-10574	Sequence 10574, A
24	92.6	7.8	1373	16	US-10-424-599-31364	Sequence 31364, A
25	89.4	7.5	650	17	US-10-767-701-4045	Sequence 4045, Ap
26	85	7.1	327	16	US-10-424-599-121152	Sequence 121152, A
27	81.4	6.8	1355	16	US-10-424-599-31363	Sequence 31363, A
28	80.8	6.8	994	16	US-10-424-599-73501	Sequence 73501, A
29	76.8	6.4	716	16	US-10-424-599-61310	Sequence 61310, A
30	67.2	5.6	1173	17	US-10-437-963-57539	Sequence 57539, A
31	62.6	5.2	774	9	US-09-770-445-895	Sequence 895, App
32	58.2	4.9	789	17	US-10-437-963-72452	Sequence 72452, A
33	55.8	4.7	308	16	US-10-424-599-121805	Sequence 121805, A
34	55.2	4.6	1417	16	US-10-424-599-97043	Sequence 97043, A
35	52.2	4.4	1296	16	US-10-425-114-13687	Sequence 13687, A
36	49.6	4.2	737	16	US-10-425-114-24295	Sequence 24295, A
37	49	4.1	1177	17	US-10-437-963-91425	Sequence 91425, A
38	43.6	3.7	383	17	US-10-021-323-2053	Sequence 2053, Ap
39	42.8	3.6	613	17	US-10-767-701-4026	Sequence 4026, Ap
40	41.2	3.5	478	18	US-10-674-124A-2371	Sequence 2371, Ap
41	41.2	3.5	867	14	US-10-123-155-20	Sequence 20, Appl
42	41.2	3.5	867	15	US-10-146-731-20	Sequence 20, Appl
43	41.2	3.5	867	15	US-10-140-472-20	Sequence 20, Appl
44	41.2	3.5	867	15	US-10-141-761-20	Sequence 20, Appl
45	41.2	3.5	867	15	US-10-142-885-20	Sequence 20, Appl

ALIGNMENTS

RESULT 1  
US-09-938-842A-2212  
; Sequence 2212, Application US/09938842A  
; Patent No. US20020160378A1  
; GENERAL INFORMATION:  
; APPLICANT: Harper, Jeff  
; APPLICANT: Kreps, Joel  
; APPLICANT: Wang, Xun  
; APPLICANT: Zhu, Tong  
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
; FILE OF INVENTION: SAME, AND METHODS OF USE  
; FILE REFERENCE: SCRIPI300-3  
; CURRENT APPLICATION NUMBER: US/09/938,842A  
; CURRENT FILING DATE: 2001-08-24  
; PRIOR APPLICATION NUMBER: US 60/227,866  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: US 60/264,647  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/300,111  
; PRIOR FILING DATE: 2001-06-22  
; NUMBER OF SEQ ID NOS: 5379  
; SEQ ID NO 2212  
; LENGTH: 3387  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-09-938-842A-2212

Query Match 85.1%; Score 1015.8; DB 9; Length 3387;  
Best Local Similarity 95.1%; Pred. No. 4.6e-304;  
Matches 1081; Conservative 0; Mismatches 2; Indels 54; Gaps 1;

QY 38 GTGACCAAGCTTAGAAGCAACCTTAGATCATGAGTGTGTAATCATCATCATCAT 97  
DB 1205 GTGACCAAGCTTAGAAGCAACCTTAGATCATGAGTGTGTAATCATCATCATCAT 1264

```
Qy 98 CAGCTGTGCTCAAAACGAGAACTATAAAAGTGGCTTCGTCTCCATATACGTAATCT 157
Db |||||
Qy 1265 CAGCTGTGCTCAAAACGAGAACTATAAAAGTGGCTTCGTCTCCATATACGTAATCT 1324
Db |||||
Qy 158 TTGTCTCTTTTGGCCAGCAGCTAGCTACAAATTTCTGGGTAGATTGTACTATGAAATGGAG 217
Db |||||
Qy 1325 TTGTCTCTTTTGGCCAGCAGCTAGCTACAAATTTCTGGGTAGATTGTACTATGAAATGGAG 1384
Db |||||
Qy 218 GAAATAGACATATGTGGTAACATTTCTTCAACTCAITTTGGCTTCCTGTACTGTTCTGT 277
Db |||||
Qy 1385 GAAATAGACATATGTGGTAACATTTCTTCAACTCAITTTGGCTTCCTGTACTGTTCTGT 1444
Db |||||
Qy 278 TCCGCTCTTTTCTCGAATCAGGCAACCAAAATCAACAGATACAAATTTTCAGTCAGTCCC 337
Db |||||
Qy 1445 TCCGCTCTTTTCTCGAATCAGGCAACCAAAATCAACAGATACAAATTTTCAGTCAGTCCC 1504
Db |||||
Qy 338 CTTCCTTCCACCACTTCATCGGTTTACTTTGTGCACCTGGACTGCTAGTGTCCGCTTATG 397
Db |||||
Qy 1505 CTTCCTTCCACCACTTCATCGGTTTACTTTGTGCACCTGGACTGCTAGTGTCCGCTTATG 1564
Db |||||
Qy 398 CTATTATTGTGTGAGTAGGGTTGCTCTACTTACCAAGTCTCTACTTTCTCCCTCATCTTGG 457
Db |||||
Qy 1565 CTATTATTGTGTGCA----- 1578
Db |||||
Qy 458 CCTCACAGTTGGCTTCACTGCTTTTCTCATATTTCTTAACTCGCAAAAGTTCACTC 517
Db |||||
Qy 1579 -----TTGGCCCTTCACTGCTTTTCTCATATTTCTTAACTCGCAAAAGTTCACTC 1630
Db |||||
Qy 518 CTTTGATAGTCAGTTCTTTTGTCTTCTCCCTCACTGTATCTCTGCTCTTCTTGTGGTCAACA 577
Db |||||
Qy 1631 CTTTGATAGTCAGTTCTTTTGTCTTCTCCCTCACTGTATCTCTGCTCTTCTTGTGGTCAACA 1690
Db |||||
Qy 578 CTGATTTCAGAAAACCTCAACTAATGATCTTAGAGTACAGTATGTGATCGGGTTCAATGTA 637
Db |||||
Qy 1691 CTGATTTCAGAAAACCTCAACTAATGATCTTAGAGTACAGTATGTGATCGGGTTCAATGTA 1750
Db |||||
Qy 638 CCATCGGTGCTTCCGCTCGGATGAGCTGTTACTATCTCTGATACAAATGCTCTTCAGGA 697
Db |||||
Qy 1751 CCATCGGTGCTTCCGCTCGGATGAGCTGTTACTATCTCTGATACAAATGCTCTTCAGGA 1810
Db |||||
Qy 698 AAGTTTTTCAGGAAGCATATCTCTCAGCAGTCACGGACTTGGCCATTTTACCAGTCTCTAG 757
Db |||||
Qy 1811 AAGTTTTTCAGGAAGCATATCTCTCAGCAGTCACGGACTTGGCCATTTTACCAGTCTCTAG 1870
Db |||||
Qy 758 TTGCAGATTGTAGTTCTCATAGACATTTTGGAAAGTGGAGTGGGAAACTTTGGCAA 817
Db |||||
Qy 1871 TTGCAGATTGTAGTTCTCATAGACATTTTGGAAAGTGGAGTGGGAAACTTTGGCAA 1930
Db |||||
Qy 818 GTGAGATCAGAACTACAACTCGGGAAGTGTATATGTTTGAATTTAGCCTTCGGCAG 877
Db |||||
Qy 1931 GTGAGATCAGAACTACAACTCGGGAAGTGTATATGTTTGAATTTAGCCTTCGGCAG 1990
Db |||||
Qy 878 CTATTTCCTGGCAAGTCTACACTCTCTGCTTGTGGGATTTGATCTTCAGTCACTCTCTG 937
Db |||||
Qy 1991 CTATTTCCTGGCAAGTCTACACTCTCTGCTTGTGGGATTTGATCTTCAGTCACTCTCTG 2050
Db |||||
Qy 938 TGTCTCCAAATTCATTAACAGCTGTGGGATTTGCCCTATATAGTTCCAGTTCGGCAGTGATAG 997
Db |||||
Qy 2051 TGTCTCCAAATTCATTAACAGCTGTGGGATTTGCCCTATATAGTTCCAGTTCGGCAGTGATAG 2110
Db |||||
Qy 998 TTTTCCATGATAGATGAGCGCATCCAAATCTCTCCATTTATTTTGTAGCTATCTCGGCT 1057
Db |||||
Qy 2111 TTTTCCATGATAGATGAGCGCATCCAAATCTCTCCATTTATTTTGTAGCTATCTCGGCT 2170
Db |||||
Qy 1058 TCCTTTTCAATTCGCTATCAGCACTACCTCGAGAAAAGAGTTGAATTAAGTCAAGCAACAA 1117
Db |||||
Qy 2171 TCCTTTTCAATTCGCTATCAGCACTACCTCGAGAAAAGAGTTGAATTAAGTCAAGCAACAA 2230
Db |||||
Qy 1118 GTGCTGTAGGAGATCTTCACTTACCTGTGTAGGAAAGGTACACAAACATACAAAGTG 1174
Db |||||
Qy 2231 GTGCTGTAGGAGATCTTCACTTACCTGTGTAGGAAAGGTACACAAACATACAAAGTG 2287
Db |||||
```

```
RESULT 2
US-09-938-842A-2212
; Sequence 2212, Application US/09938842A
; Publication No. US20040009476A9
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Krepes, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIP1300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 2212
; LENGTH: 3387
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-2212

Query Match 85.1%; Score 1015.8; DB 11; Length 3387;
Best Local Similarity 95.1%; Pred. No. 4.6e-304;
Matches 1081; Conservative 0; Mismatches 2; Indels 54; Gaps 1;

Qy 38 GTGACCGAAGCTTAGAAGCAACCTTATAGATCATGATGAGTGGTAACGAATCATCATCAT 97
Db |||||
Qy 1205 GTGACCGAAGCTTAGAAGCAACCTTATAGATCATGATGAGTGGTAACGAATCATCATCAT 1264
Db |||||
Qy 98 CAGCTGTGCTCTCAAAACGAGAACTATAAAAGTGGCTTCGTGTCCATATACGTAATCT 157
Db |||||
Qy 1265 CAGCTGTGCTCTCAAAACGAGAACTATAAAAGTGGCTTCGTGTCCATATACGTAATCT 1324
Db |||||
Qy 158 TTGTCTCTTTTGGCCAGCAGCTAGCTACAAATTTCTGGGTAGATTGTACTATGAAATGGAG 217
Db |||||
Qy 1325 TTGTCTCTTTTGGCCAGCAGCTAGCTACAAATTTCTGGGTAGATTGTACTATGAAATGGAG 1384
Db |||||
Qy 218 GAAATAGACATATGTGGTAACATTTCTTCAACTCAITTTGGCTTCCTGTACTGTTCTGT 277
Db |||||
Qy 1385 GAAATAGACATATGTGGTAACATTTCTTCAACTCAITTTGGCTTCCTGTACTGTTCTGT 1444
Db |||||
Qy 278 TCCGCTCTTTTCTCGAATCAGGCAACCCAAATCAACAGATACAAATTTTCAGTCAGTCCC 337
Db |||||
Qy 1445 TCCGCTCTTTTCTCGAATCAGGCAACCCAAATCAACAGATACAAATTTTCAGTCAGTCCC 1504
Db |||||
Qy 338 CTTCCTTCCACCACTTCATCGGTTTACTTTGTGCACCTGGACTGCTAGTGTCCGCTTATG 397
Db |||||
Qy 1505 CTTCCTTCCACCACTTCATCGGTTTACTTTGTGCACCTGGACTGCTAGTGTCCGCTTATG 1564
Db |||||
Qy 398 CTATTATTGTGTGAGTAGGGTTGCTCTACTTACCAAGTCTCTACTTTCTCCCTCATCTTGG 457
Db |||||
Qy 1565 CTATTATTGTGTGCA----- 1578
Db |||||
Qy 458 CCTCACAGTTGGCCCTTCACTGCTTTTCTCATATTTCTTAACTCGCAAAAGTTCACTC 517
Db |||||
Qy 1579 -----TTGGCCCTTCACTGCTTTTCTCATATTTCTTAACTCGCAAAAGTTCACTC 1630
Db |||||
Qy 518 CTTTGATAGTCAGTTCTTTTGTCTTCTCCCTCACTGTATCTCTGCTCTTCTTGTGGTCAACA 577
Db |||||
Qy 1631 CTTTGATAGTCAGTTCTTTTGTCTTCTCCCTCACTGTATCTCTGCTCTTCTTGTGGTCAACA 1690
Db |||||
Qy 578 CTGATTTCAGAAAACCTCAACTAATGATCTTAGAGTACAGTATGTGATCGGGTTCAATGTA 637
Db |||||
Qy 1691 CTGATTTCAGAAAACCTCAACTAATGATCTTAGAGTACAGTATGTGATCGGGTTCAATGTA 1750
Db |||||
Qy 638 CCATCGGTGCTTCCGCTCGGATGAGCTGTTACTATCTCTGATACAAATGCTCTTCAGGA 697
Db |||||
Qy 1751 CCATCGGTGCTTCCGCTCGGATGAGCTGTTACTATCTCTGATACAAATGCTCTTCAGGA 1810
Db |||||
Qy 698 AAGTTTTTCAGGAAGCATATCTCTCAGCAGTCACGGACTTGGCCATTTTACCAGTCTCTAG 757
Db |||||
Qy 1811 AAGTTTTTCAGGAAGCATATCTCTCAGCAGTCACGGACTTGGCCATTTTACCAGTCTCTAG 1870
Db |||||
Qy 758 TTGCAGATTGTAGTTCTCATAGACATTTTGGAAAGTGGAGTGGGAAACTTTGGCAA 817
Db |||||
Qy 1871 TTGCAGATTGTAGTTCTCATAGACATTTTGGAAAGTGGAGTGGGAAACTTTGGCAA 1930
Db |||||
Qy 818 GTGAGATCAGAACTACAACTCGGGAAGTGTATATGTTTGAATTTAGCCTTCGGCAG 877
Db |||||
Qy 1931 GTGAGATCAGAACTACAACTCGGGAAGTGTATATGTTTGAATTTAGCCTTCGGCAG 1990
Db |||||
Qy 878 CTATTTCCTGGCAAGTCTACACTCTCTGCTTGTGGGATTTGATCTTCAGTCACTCTCTG 937
Db |||||
Qy 1991 CTATTTCCTGGCAAGTCTACACTCTCTGCTTGTGGGATTTGATCTTCAGTCACTCTCTG 2050
Db |||||
Qy 938 TGTCTCCAAATTCATTAACAGCTGTGGGATTTGCCCTATATAGTTCCAGTTCGGCAGTGATAG 997
Db |||||
Qy 2051 TGTCTCCAAATTCATTAACAGCTGTGGGATTTGCCCTATATAGTTCCAGTTCGGCAGTGATAG 2110
Db |||||
Qy 998 TTTTCCATGATAGATGAGCGCATCCAAATCTCTCCATTTATTTTGTAGCTATCTCGGCT 1057
Db |||||
Qy 2111 TTTTCCATGATAGATGAGCGCATCCAAATCTCTCCATTTATTTTGTAGCTATCTCGGCT 2170
Db |||||
Qy 1058 TCCTTTTCAATTCGCTATCAGCACTACCTCGAGAAAAGAGTTGAATTAAGTCAAGCAACAA 1117
Db |||||
Qy 2171 TCCTTTTCAATTCGCTATCAGCACTACCTCGAGAAAAGAGTTGAATTAAGTCAAGCAACAA 2230
Db |||||
Qy 1118 GTGCTGTAGGAGATCTTCACTTACCTGTGTAGGAAAGGTACACAAACATACAAAGTG 1174
Db |||||
Qy 2231 GTGCTGTAGGAGATCTTCACTTACCTGTGTAGGAAAGGTACACAAACATACAAAGTG 2287
Db |||||
```



Db 1751 CCATCGGTGCTCCGCTGGGATTTGGACTGTATTCTATCTCTGATACAAATGCTCTTCAGGA 1810  
QY 698 AAGTTTTTCAGGAAGCATACATCTCAGCAGCTACGAGCTTTGGCCATTTTACCAGTCTCTAG 757  
Db 1811 AAGTTTTTCAGGAAGCATACATCTCAGCAGCTACGAGCTTTGGCCATTTTACCAGTCTCTAG 1870  
QY 758 TTCCGAGTTGTGTAGTTCTCATAGGACTTTTTCGAAAGTGGAGAGTGGGAAACTTTTGCCAA 817  
Db 1871 TTCCGAGTTGTGTAGTTCTCATAGGACTTTTTCGAAAGTGGAGAGTGGGAAACTTTTGCCAA 1930  
QY 818 GTGAGATGAGAACTACAACTCGGAAAGTGTTCATATGTTTGTAGCTTTAGCTCGGCAG 877  
Db 1931 GTGAGATGAGAACTACAACTCGGAAAGTGTTCATATGTTTGTAGCTTTAGCTCGGCAG 1990  
QY 878 CTATTTTCCTGGCAAGTCTACACTCCTCGTCTTGTGGGATTTGATCTTCGAGTCACTCTCTG 937  
Db 1991 CTATTTTCCTGGCAAGTCTACACTCCTCGTCTTGTGGGATTTGATCTTCGAGTCACTCTCTG 2050  
QY 938 TGTTCCTCAATTCATAACAGCTGTGGGATTTGCTATAGTTCCAGTTGGCGAGTGATAG 997  
Db 2051 TGTTCCTCAATTCATAACAGCTGTGGGATTTGCTATAGTTCCAGTTGGCGAGTGATAG 2110  
QY 998 TTTTCCATGATGAATGGAGCGCATCCAAATCTTCTCCATTATTTTAGCTATCTCGGCT 1057  
Db 2111 TTTTCCATGATGAATGGAGCGCATCCAAATCTTCTCCATTATTTTAGCTATCTCGGCT 2170  
QY 1058 TCCTTTTCATTTCGCTATCAGCACTACCTCGACGAAAGAGTTGAATCTAGCCACACAA 1117  
Db 2171 TCCTTTTCATTTCGCTATCAGCACTACCTCGACGAAAGAGTTGAATCTAGCCACACAA 2230  
QY 1118 GTGCTGTAGGAGATCTTCATCTACCTGTTGAGGAAAGTGCACACAAATACAAAGTG 1174  
Db 2231 GTGCTGTAGGAGATCTTCATCTACCTGTTGAGGAAAGTGCACACAAATACAAAGTG 2287

RESULT 3

US-09-770-152-44  
; Sequence 44, Application US/09770152  
; Publication No. US20020040489A1  
; GENERAL INFORMATION:  
; APPLICANT: Gorchach, Jorn  
; APPLICANT: An, Yong-Qiang  
; APPLICANT: Hamilton, Carol M.  
; APPLICANT: Price, Jennifer L.  
; APPLICANT: Raines, Tracy M.  
; APPLICANT: Yu, Yang  
; APPLICANT: Rameaka, Joshua G.  
; APPLICANT: Page, Amy  
; APPLICANT: Matthew, Abraham V.  
; APPLICANT: Ledford, Brooke L.  
; APPLICANT: Woessner, Jeffrey P.  
; APPLICANT: Haas, William David  
; APPLICANT: Garcia, Carlos A.  
; APPLICANT: Kricker, Maja  
; APPLICANT: Slader, Ted  
; APPLICANT: Davis, Keith R.  
; APPLICANT: Allen, Keith  
; APPLICANT: Hoffman, Neil  
; APPLICANT: Hurban, Patrick  
; TITLE OF INVENTION: Expressed Sequences of Arabidopsis  
; FILE OF INVENTION: thaliana  
; FILE REFERENCE: 2025US (PARA-014PRV)  
; CURRENT APPLICATION NUMBER: US/09/770,152  
; CURRENT FILING DATE: 2001-01-26  
; PRIOR APPLICATION NUMBER: 60/178,503  
; PRIOR FILING DATE: 2000-01-27  
; NUMBER OF SEQ ID NOS: 999  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 44  
; LENGTH: 592  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-09-770-152-44

Query Match 38.5%; Score 459.8; DB 9; Length 592;  
Best Local Similarity 91.6%; Pred. No. 9.4e-132;  
Matches 501; Conservative 0; Mismatches 37; Indels 9; Gaps 1;  
QY 37 GGTGACCAAGAACTTAAAGCAAACTTATAGATCATAGAGTGGTAACTCAATCATCATCA 96  
Db 55 GGTGACCAAGAACTTAAAGCAAACTTATAGATCATAGAGTGGTAACTCAATCATCATCA 105  
QY 97 TCAGCTGTGCTCAAAACCGAGAACTATAAAGTGGCTTCGTTGCTCCATATACGTAATC 156  
Db 106 TTTTCAGTACCTCAAAACCGAGAACTGTAGAGTGGCTCCGTTCTCCATATACGTAATC 165  
QY 157 TTTTGTCTCTTTTGGCAGCCACTAGCTACAAATCTCTGGGTAGATTTGATCTATGAAATGA 216  
Db 166 TTTTGTCACTCTTCTGCCAACCACTTGTCTACAGTTCTGGGTAGACTGTACTATGAAATGA 225  
QY 217 GGAATAGACATATGTGTAAACACTTCTTCAACTCATTTGGCTTCCCTGTACTGTCTTG 276  
Db 226 GGGAAAAGCACATATGTGTAAACACTTCTTCAACTCATTTGGCTTCCCTGTACTGTCTC 285  
QY 277 TTTCCGCTCTTTTCTCGAATCAGGCAACCCAAATCAACAGATACAAATTTTCAGTCAGTCC 336  
Db 286 TTTCCGCTCTTTTCTCGAATCAGGCAACCCAAATCAACAGATACAAATTTTCAGTCAGTCC 345  
QY 337 CCTTCTTCAACCCCTTGTGCTTGTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 396  
Db 346 CCTTCTTCAACCCCTTGTGCTTGTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 405  
QY 397 GCTTATTTGTCTGAGTAGGGTGTCTTCACTTACAGTCTTCTTCTTCTTCTTCTTCTTCTTCT 456  
Db 406 GCTTATTTGTCTGAGTAGGGTGTCTTCACTTACAGTCTTCTTCTTCTTCTTCTTCTTCTTCT 465  
QY 457 GCTTCACTGTTGGCTTCACTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 516  
Db 466 GCTTCACTGTTGGCTTCACTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 525  
QY 517 CCTTGTAGTAGTCACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 576  
Db 526 CCTTGTAGTAGTCACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 585  
QY 577 ACTGATT 583  
Db 586 ACTGATT 592

RESULT 4

US-10-424-599-2524  
; Sequence 2524, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J  
; APPLICANT: Kovalic, David K  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE OF INVENTION: plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 2524  
; LENGTH: 1294  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_102285C.1  
US-10-424-599-2524

Query Match 29.3%; Score 349.8; DB 16; Length 1294;  
Best Local Similarity 59.5%; Pred. No. 2.2e-97;  
Matches 622; Conservative 0; Mismatches 402; Indels 21; Gaps 1;



Db 1020 ATCAATTCGTGCTCTGGTCTTCTGTCTATAGCCCTATTTTGCAGTGGGGAATGGCGTAC 1079  
Qy 810 TTTCGCAAGTGAAGTGAAGAACTACAACTCGGAAAGTGTCTATATGTTTTCGACTTTAGC 869  
Db 1080 TTTCGCAAGTGAAGTGAAGAACTACAACTCGGAAAGTGTCTATATGTTTTCGACTTTAGC 1139  
Qy 870 CTCGGCAGCTATTTCTGCGAAGTCTACACTCTCTGCTCTGTTGGGATTTGATCTTCAGTTC 929  
Db 1140 TTGGACTTCAATAGCTCGCAGGATGCTCTGTTGGTGTGTTGGCTTGTATCTTCTAGT 1199  
Qy 930 ATCTCTGTGTTCTCCAAATTCATAACAGCTGTGGGATTCCTATATAGTTTCCAGTTTCGGC 989  
Db 1200 GTCTCTCTCTACTCCAAATGTTATAAGCAAGTTCTTTAGCGGTAACCTCTATTTGCTGC 1259  
Qy 990 AGTGATAGTTTTCATATGATAGTACGACGATCCAAATCTTCTCCATTAATTTTAGCTAT 1049  
Db 1260 TGTATAGTTTTCATGATAAGATGAATGGGGTGAAGATAATTTCTATGCTTTTGGCTCT 1319  
Qy 1050 CTGCGGCTTCTTTCATTCGTTCTATCAGCACTACCTCGAGCAAGAAAG 1098  
Db 1320 ATGGGGTTTTGCTCTTATATTTATCAGAATTTATCTTTGATGATTTCAAAG 1368

## RESULT 6

US-10-437-963-82061/c

; Sequence 82061, Application US/10437963

; Publication No. US20040123343A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; APPLICANT: Wu, Wei

; APPLICANT: Boukharov, Andrey A.

; APPLICANT: Barbazuk, Brad

; APPLICANT: Li, Ping

; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with

; FILE REFERENCE: 38-21(53221)B

; CURRENT APPLICATION NUMBER: US/10/437,963

; CURRENT FILING DATE: 2003-05-14

; NUMBER OF SEQ ID NOS: 204966

; SEQ ID NO 82061

; LENGTH: 1513

; TYPE: DNA

; ORGANISM: Oryza sativa

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_81525C.1

US-10-437-963-82061

Query Match 23.4%; Score 279; DB 17; Length 1513;  
Best Local Similarity 55.8%; Pred. No. 2.4e-75;  
Matches 556; Conservative 0; Mismatches 435; Indels 6; Gaps 1;

Qy 107 CTCAAACCGAGAACTATAAAGTGGCTTCTCCATATACGTAATCTTTTGGCTCT 166  
Db 1278 CTCCTGTCAAGAGTTGGCAATGGTCTGATGGTGGAGTTAAATGTTCTTCTCATTTG 1219  
Qy 167 TTTCGCGCCACTAGTCAATCTCGGTAGATTGCTATCAAAATGAGGAAATAGCA 226  
Db 1218 CCGGTCAACCGCATCTACACTTCTTGGGAGTTCTACTACAAATCAAGGCGCAATAGCA 1159  
Qy 227 CATATGTGTTAACTCTTCAACTCATTTGGCTTCCCTGCTACTGTTCTGTTCCGCTTCT 286  
Db 1158 AGTGGATGTCACATTCGTCGAACTGTGGCTTTCCGATTTTGTTCATTTGCTTATTC 1099  
Qy 287 TTTCGGAATCAGGCAACCAATCAACAGATACAAATTTTCAGTCAGTCCCC-----TT 340  
Db 1098 TTTCCTCATTCAAAGACATCTTCTACACAAACAGTCACTAGTAGTCTGCGCCCTACAAATTT 1039  
Qy 341 CTTTCACCAACCTTGCATCGGTTTACTTGTGCACTGGACTGTAGTTCGCTTATGCTT 400  
Db 1038 CTATCCCCAAAATTACTCTGATATATGTTGTCTCGGCGCTCATCATTTGCTGCAGACGACT 979

Qy 401 ATTTGTCTGAGTAGGTTGCTCTACTTACCAGTCTCTACTTTTCCCTCACTTTGGCCT 460  
Db 978 TGATGTAATTCCTATGCGCTACTATATCTTCGGTCTCAACATATTCGCTCACTGTGCTA 919  
Qy 461 CACAGTTGGCCTTCACTGCTTTTCTCATATTTCTTAACTCGGAAAGTTCACCTCTT 520  
Db 918 GTCACTTTGGCCTTCAATGCTGTCTCTCATATTTCTCAATGCTCAAAATTTCAACCTTC 859  
Qy 521 TGATAGTCAAGTTCTTTGCTTCTCCCTCACTGATATCTCTGCTCTTCTTGTGGTCAACACTG 580  
Db 858 TGATTTTCAATTCCTGAGTCTCTCTTACGTTTCTGCTTCACTCTCTGAGTTGATGAAG 799  
Qy 581 AFTCAGAAAACTCAACTAATGTATCTAGAGTACAGTATGTGATCGGGTTTCATATGTACCA 640  
Db 798 ATTCTCAGGAACTACTAGTATATCACATGGGAAGTACATTTTGGGTTTCTCTGTTGACAC 739  
Qy 641 TCGGTGCTTCGCTGGGATTTGAGCTGTACTATCTCTGATACAAATGCTCTTTCAGAAAG 700  
Db 738 TAGGGGCATCAGCTACATACTCGCTCATTTCTCTCCCTGATGCAAGTCAATTTGAGAAG 679  
Qy 701 TTTTCACGAAGCATACATCCCTCAGCAGTCACGGACTTGGCCATTTACCACTCTCTAGTTG 760  
Db 678 TTATTAAGAGGAGACCTTCTCAGTTGTGTTGAACATGCAGATATATACAGCTCTCGTGG 619  
Qy 761 CGAGTTGTGTAGTTCTCATAGGACTTTTTCGAAAGTGGAGAGTGGGAAACTTTGCCAAGTG 820  
Db 618 CAACATTTGGCTTCTCTGTTGGTTATTTGCAAGTGGTGAATGGATGACTTTTACAAGGAG 559  
Qy 821 AGATGAGAAACTACAAACTCGGAAAGTGTCTATGTTTGTGCTTTAGCTTTAGCCTCGGAGCTA 880  
Db 558 AGATGATGCAATTCAGTCTCGGAAAGTGTCTATATGTAATGACACTGTCTGTGGAGCGGCTA 499  
Qy 881 TTTCCTGCAAGTCTACACTCTCTGCTCTTGTGGGATGATCTTCAGTCACTCTCTGCTGT 940  
Db 498 TATCTTGCAGGTAGCATCAGTTGAGTGGTGGGATGATCTTTTGTGTTGTCATCGCTGT 439  
Qy 941 TCTCAATTTCCATAACAGCTGTGGGATTTGCCCTATAGTTTCCAGTTTCGGGAGTGATAGTTT 1000  
Db 438 TTTCAATGTGATAAGCACCCCTAGCTTACCCATCATCTCTGTTTGTGCTGATTTTCT 379  
Qy 1001 TCCATGATAGAAATGGAACGATCCAAATCTTCTCCATTAATTTTAGCTATCTGCGGCTTCC 1060  
Db 378 TTCATGCAAGATGGATGGAGTAAAGATTATAGCTATGCTGATGGCCATTTGGGATTTA 319  
Qy 1061 TTTCAATTCGTCTATCAGCACTACCTCGACGAAAGAA 1097  
Db 318 TGTCAATGGCCACCACCAATTATATGTTGATGGCAAGAA 282

## RESULT 7

US-10-424-599-75161

; Sequence 75161, Application US/10424599

; Publication No. US20040031072A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa Thomas J

; APPLICANT: Kovalic David K

; APPLICANT: Zhou Yihua

; APPLICANT: Cao Yongwei

; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with

; FILE REFERENCE: 38-21(53223)B

; CURRENT APPLICATION NUMBER: US/10/424,599

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 285684

; SEQ ID NO 75161

; LENGTH: 997

; TYPE: DNA

; ORGANISM: Glycine max

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_38884C.1

US-10-424-599-75161

Query Match		20.7%;	Score 247.2;	DB 16;	Length 997;
Best Local Similarity		57.3%;	Pred. No. 1.4e-65;		
Matches 500;		Conservative 0;	Mismatches 348;	Indels 24;	Gaps 2;
Qy	118	AACTATAAAGGTGGCTTCGTGTCCTCCATATACGTAATCTTTGTGCTCTTTTGGCAGCCA	177		
Db	125	AAATACAAATGGTGGCTACGGTGTGCTCTACATAATCTTCTCTGTAGGCCAATCT	184		
Qy	178	CTAGCTACAAATCTGGGTAGATTGTACTATGAATAATGGAGAAATAGCACATATGTGGTA	237		
Db	185	CGAGCACTCTTTGGGAAGATTATATATGACAAATGGTGGTAAATAGCAAAATGGATGGCA	244		
Qy	238	ACACTTCTTCAACTCATTTGCTGCTCCCTGACTGTGTTGTCGGTCTCTTTCTCGA---	294		
Db	245	ACATTGTTCATATCAGCTGGAAATCCCTGTACTTCTTCCACCTCTTTTACTTCCCAAGA	304		
Qy	295	-----ATCAGGCAACCCCAATCAACAGATACAAAATTTTCAGTCAGTCC	336		
Db	305	CAAAACATGCCAAGTTTAATAACAAATCCAGCAATATGATTATTCACAAAACAAA	364		
Qy	337	CCTTCCTTCAACACCTTGGCATCGGTTTACTTGTGCACTGGACTGTAGTGTCCGCTTAT	396		
Db	365	CCAAAATTTCTCCACCTCTGTTTCTCTATCTAGCTTTTGGCCTAAATTTCTCACAGGGAC	424		
Qy	397	GCTTATTTGTCTGCAGTAGGGTTGCTCTACTTACCAGTCTACTTTCTCCCTCATCTTG	456		
Db	425	AACTTGATGATTTCTTATGGACTTTTATATCTCCCTCTTCCACCTATTTCTTACTATGT	484		
Qy	457	GCCTCACAGTTGGCCCTTCACTGCCTTTTCTCATATTTCTTAACTCGCAAAAGTTCACT	516		
Db	485	GCCACCAATAGTCTTCAACGGGTGTCCTTCTTCTTCTCAATTCACGAAATTCACA	544		
Qy	517	CCTTTGATGACAGTTCTTGGCTTCTCCTACTGATATCCTCTGCTCTTCTTGTGGTCAAC	576		
Db	545	GCATTCAATAAATTTCTGTAGTCTCTTCTTACCATATCAGCTTCTCTGCTTGCATCAAC	604		
Qy	577	ACTGATTC--AGAAAATCAACTAAATGATCTAGAGTACAGTATGTATCGGGTTCATA	633		
Db	605	TCCGATTCGAGAGGACTCAACCGGTCTTTCCAGAGAAAGCATGTAATTTGGATTTCTC	664		
Qy	634	TGTACCATCGGTGCTTCGGCTGGGATTTGGACTGTGTACTATCTCTGATACAAATGCTTTC	693		
Db	665	TGCACCATTTGGTGCATCAGCTACATCTCTTTGTACCTCTCTCTGTGACGCTTCTTTT	724		
Qy	694	AGGAAAGTTTCAACGAGCATACATCTTCAGAGTACAGGACTTGGCCATTTACAGTCT	753		
Db	725	CAGAAAGTTATAAGAGAGAAACCTTTTCTGCTGTGTTGGACATGCAATTTTACCCATCC	784		
Qy	754	CTAGTTGCGAGTTGTGTAGTTCTCATAGGACTTTTTCGAAGTGGAGAGTGGGAACTTTG	813		
Db	785	TTCATTTGCTCATGTGCTTGTGTAGTAGGGTTGTTTTCGAAGTGGAGAGTGGGAAAGTTTG	844		
Qy	814	CCAAGTGAGATGAGAAATACAAACTCGGGAAAGTGTATATTTTGTGACTTTTAGCCCTCG	873		
Db	845	AACATGAGATGAAGGATATGATAAGGAAGTGTGCCCGTTGTAATGACTGTACTATAG	904		
Qy	874	GCAGCTATTTCTGGCAAGTCTACACTCCTGGTCTTGTGGGATTTGATCTTCGAGTCATCC	933		
Db	905	ATTGCTGTGACATGTATATATGCTCACTAGGGATGCTAGGGGGACGTTTGGAGAGTCT	964		
Qy	934	TCTGTGTTCTCCAATTTCCATAACAGCTGTGGG	965		
Db	965	GCAATTTGTCGGATGTGAGAAGTACCTGGGG	996		

RESULT 8

US-10-437-963-43860/c  
; Sequence 43860, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 43860  
; LENGTH: 2381  
; TYPE: DNA  
; ORGANISM: Oryza sativa  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(2381)  
; OTHER INFORMATION: unsure at all n locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_46977C.1  
; US-10-437-963-43860

Query Match

Best Local Similarity 20.5%; Score 244.2; DB 17; Length 2381;  
Matches 544; Conservative 0; Mismatches 458; Indels 15; Gaps 1;

Qy	89	CATCATCATCAGCTGTGCCCTCAAAACGAGAACTATAAAAGTGGCTTCGTGTCTCCATAT	148
Db	2284	CAGCAGCTCGGAAACACACGAGGCAAGCAATTCGATGTGGCGCATGTGGCAGTGG	2225
Qy	149	ACGTAATCTTTGCTCTTTTGGCAGCCACTAGTACAAATCTTGGGTAGATTGTACTATG	208
Db	2224	ACGCTCTTCTCTCATCGCGGTGAGACATCGGCGACACTGCTGGGAGATCTACTACA	2165
Qy	209	AAATGAGGAAATAGCACATATGTGTAACACTTCTTCAACTCATTTGGCTTCCCTGTAC	268
Db	2164	CCCAAGCGCGCAGCAAGTGGATATCGGGTTCGTGCGGACGGCCGCTTCCCGATAC	2105
Qy	269	TGGTCTTGTTCGGCTCTTTTCGAATCAGGCAACCCAAATCAACAGATACAAATTTCA	328
Db	2104	TGTTCTTTCACCTCTTCTTCTTCCCTCGAAGTCACCGTCTC-----CT	2060
Qy	329	GTAGTCCCTTCTTTCACCACTTCGATCGGTTTACTTGTGCACTGGAGTGTAGTGT	388
Db	2059	GCACCAACATCTCCCATGGCAAGCTCGCGGTGATATACATCGTCTCGGCTCATATCG	2000
Qy	389	CCGCTTATGCTTATTTCTGTCAGTAGGGTTGCTCTACTTACCAGTCTCTACTTTCTCCC	448
Db	1999	CCGCGAGCAGATGATGTACACTGGTGGCTCAAGTACCTCCCGCTTCGACCTACTCGC	1940
Qy	449	TCATCTTGGCTCACAGTTGGCTTCACTGCTTTTTTCTCATATTTCTTAACTTCGCAAA	508
Db	1939	TCATCTGCGCAGCAGCTCGCGTTCAATGTGCTTCTCTCATAGTGTCTCACTCCAGA	1880
Qy	509	AGTTCACTCTTTGATAGTGTCTTTTGTCTTCTCCTCACTGTATCTCTGCTCTCTTG	568
Db	1879	AGGTCATCTCTCTGATATTCAACTCCGTGCTGTGCTCACCATGCTCGCTTCGCTCATCG	1820
Qy	569	TGGTCAACACTGATTCAGAAAACTCAACTAATGTATCTAGAGTACAGTATGTATCGGGT	628
Db	1819	GAGTCAGCAAGGAGTCTCAGGGGGTCAACCGCGTCTCGGGAGGGAAGTATCTGCTCGTT	1760
Qy	629	TCATATGTACATCGGTGCTTCGCTGGGATTTGAGCTGTTACTATCTCTGATACAAATGC	688
Db	1759	TCGTGCTGACGCTGGGGCGTCTGTGCACCTACTCGCTGATCTCTCGCTCATGACGTCA	1700
Qy	689	TCCTCAGGAAAGTTTTCACGAAGCATACATCTCAGCAGTACCGACTTGGCCATTACCT	748
Db	1699	CCCTCGAGACCATCATCAAGAAGCACCTTCTCAGCCGCTCTCAACATGCAGATCTACA	1640
Qy	749	AGTCTCTAGTTGCGAGTTGTGTAGTTCTCATAGGACTTTTGTGCAAGTGGAGAGTGGAAA	808

Db 1639 CGCGGCTGCTGGCGAAGCGCGCTCGGTGTCGGGCTGTCGCGAGCGGCGAGTGAGGT 1580  
QY 809 CTTTGGCAAGTGAGATGAGAAATACAAACTCGGGAAGTGTATATGTTTGGCTTTAG 868  
Db 1579 CGGTGAGGGGGAGATGAACGGCTTCAGGTTCGGGCGAGTTCCTACCTGATGACGCTGC 1520  
QY 869 CTTGGCAGCTATTTCTGCGCAAGTCTACACTCCTGGTCTGTGGGATGATCTTCGAGT 928  
Db 1519 TGTGGGCGGCGGTGCTGTCGCGAGGTGGCCAACTCGGGGTGCTCGGCCTCATCTTCGAGG 1460  
QY 929 CATCTCTGTGTTCTCCAAATCCATAACAGCTGTGGATTCGCTATAGTTTCCAGTTGCGG 988  
Db 1459 TGTGGGCGCTCTCTCCAGCTGATCAGACGGTGTGTCGCGGCTATCCGCTTCTTCG 1400  
QY 989 CAGTGATGATTTTCCATGATAGATGAGCAATCCAAATCTTCTCCATTTATTTAGCTA 1048  
Db 1399 CGGTGCTGTGTTTCCAGCAGGATGAACGGGGTGAAGATCGTGGCCATGCTGATTCGAA 1340  
QY 1049 TCTGGCGCTTCTTTTCAATGCTATCAGCACTACCTCGACGAAAGAAAGTTGAATA 1105  
Db 1339 TTTGGGGATTTATTTCTGTATCTGTTTCTGACCACTATCTAGATGGCAAGAAAGCAAGA 1283

RESULT 9

US-10-416-898-9  
; Sequence 9, Application US/10416898  
; Publication No. US20040172670A1  
; GENERAL INFORMATION:  
; APPLICANT: Yale University  
; APPLICANT: Walker, Elsbeth  
; APPLICANT: Dellaportia, Stephen  
; TITLE OF INVENTION: MAIZE YELLOW STRIPE1 AND RELATED GENES  
; FILE REFERENCE: 44574-5106-US  
; CURRENT APPLICATION NUMBER: US/10/416,898  
; CURRENT FILING DATE: 2003-05-16  
; PRIOR APPLICATION NUMBER: PCT/US01/43101  
; PRIOR FILING DATE: 2001-11-16  
; PRIOR APPLICATION NUMBER: US 60/249,222  
; PRIOR FILING DATE: 2000-11-16  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 9  
; LENGTH: 83698  
; TYPE: DNA  
; ORGANISM: Zea mays  
US-10-416-898-9

Query Match 18.0%; Score 215.4; DB 17; Length 83698;  
Best Local Similarity 52.2%; Pred. No. 1.3e-54;  
Matches 508; Conservative 0; Mismatches 456; Indels 9; Gaps 1;  
QY 125 AAAGGTGGCTTGTGTCTCCATATAGTAATCTTTGCTCTCTTTGCGAGCCACTAGCTA 184  
Db 67226 AATGGTGGATCTGTGTTTCAATAGCACTCTCTTCTCACTCTGCTCAAGCAATTCG 67285  
QY 185 CAATTCGGGTAGATGTACTATGAAATCGAGAAATAGCACATATGTGTAACACTTC 244  
Db 67286 TTCTTCTGTCGTTTTTATTACAAGAGGTGGAACAGTAATGATCTACTCTTG 67345  
QY 245 TTCAACTCATTTGCTTCCCTGACTGTTGCTGTTCCGCTTCTTTTCTCGAAATCAGGCAAC 304  
Db 67346 TTCAAACTGGTGGCTTCCCAATCTTTATCTCCCTCTTCTCTCTCTCTCTTCAAT 67405  
QY 305 CCAAAATCAACAGATCAAAATTTCAAGTCAGTCAGTCCTTCTTCCACCCTTGCATCGTTT 364  
Db 67406 CTTCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 67456  
QY 365 ACTTGTGCACTGACTGTAGTGTGCGCTTATGCTTATTTGCTGAGTGGTGTCTCT 424  
Db 67457 ATCTTCTCTGTTGTTGCTATTTGTTAGTAATTTTATACCTCTGTGACTTTGT 67516  
QY 425 ACTTACCACTGTCTACTTCTCTCCCTCATCTTGGCCTCAAGTTGGCTTCACTGCTTTT 484

Db 67517 ATCTCTCTCTTCAACTTATTCGATTTCTATGTGCTTTCACAGTTAGCTTTCAATGTTGTTCT 67576  
QY 485 TCTCATATTTCTTAACTCGCAAAAGTTTCACTCTCTTTGATAGTACAGTCTCTTCTCTCTCC 544  
Db 67577 TCTATTTATGATCAATCTCTCAGAAATCACTGTTTGAATTTCTCTCAGTGTGTTTCT 67636  
QY 545 TCACGTGATCTCTGCTCTTCTTGTGGTCAACACTGATTCAGAAACTCAACTPAATGTAT 604  
Db 67637 TCTCTATCTCTGCTGTGTTGTTTCTTTCAGATGATTTCAATAGCCCATCAGGAGATT 67696  
QY 605 CTAGAGTACAGTATGATCGGGTTCATATGTACCATCGTCTCTCCGCTGGATGGAC 664  
Db 67697 CTAAGTGGAGTTACTTGTGTTGTTCTGTCGAGTGTCTTCTCTCTTCTTCTTCTCTCT 67756  
QY 665 TGTACTATCTCTGATACAAATGCTCTTCAGGAAAGTCTTTCAGAGAGCATACATCCTCAG 724  
Db 67757 TTCAGTCTCTCTTATGACAGTCTTCTTTCAGAGAGTCTCAAGAGTGAAGTCTCTCTCTA 67816  
QY 725 CAGTCAAGGACTTGGCCATTTTACCAGTCTCTTGTGCGAGTGTGTGTAGTTTCTCATAGGAC 784  
Db 67817 TGGTCTCTGAGATGCAAAATCTATACGTGCTTGTGCTTCTTGTGTAGCGGTTATCGGAT 67876  
QY 785 TTTTTCAGAGTGGAGTGGGAAACTTTGCCAAGTGAGATGAGAACTACAACTCGGA 844  
Db 67877 TGTTCGCAAGCGGGAATGGAATGTTGTTGAGTGTGAGATGGAAGAGTTTCAGGAAGGTC 67936  
QY 845 AAGTGTCTATGCTTTTGTAGCTTTTGTAGCTCGGAGTATTTCTTGGCAAGTCTACACTCTCTG 904  
Db 67937 AAGTCATTTATGTTTGTAGCTTTTGTGCGGGCAGCGGTTTGTGTCATTTGGTGTGTAG 67996  
QY 905 GTCTTGTGGGATTTGATCTTCGAGTCACTCTCTGTGTTCTTCCAAATTCATAACAGCTGTGG 964  
Db 67997 GAGCGGTGTCGCTTATATTTCTGCTGTCTTCTGCTGTTTCAAACCTTATTAGTACGCTCT 68056  
QY 965 GATTCCTATAGTTCAGTTCGCGGAGTATGATTTTCCATGATAGATGGAGCATCCA 1024  
Db 68057 CACTCATTTGTTACGCTCTCTCGCGGCAATTCGCGGTGTTCCATGACAAGCTGACTGAGGTTA 68116  
QY 1025 AAATCTTCTCCATTTTATTTAGTATCTGCGGCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1084  
Db 68117 AGATGTTGGGATGCGCATCGCTTCCACAGGATTCAGTTTATATCTACCAAGACTATC 68176  
QY 1085 TCGACGAAAGAA 1097  
Db 68177 TTGATGACTTGAA 68189

RESULT 10

US-10-424-599-114931  
; Sequence 114931, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 114931  
; LENGTH: 522  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_74793C.1  
US-10-424-599-114931

Query Match 13.2%; Score 157.2; DB 16; Length 522;  
Best Local Similarity 58.0%; Pred. No. 9e-38;  
Matches 296; Conservative 0; Mismatches 213; Indels 1; Gaps 1;

472	Qy	TTCACTGCCCTTTTCTCATATTCTTAACTCGAAAGTTCACTCCTTTGTATAGTCAGT	531
6	Db	TTCAATGCAGTGTCACTTCTTCATCAATCCCAAAGTTCACATGCAATTCCTTAAAC	65
532	Qy	TCTTTGCTTCTCCTCACTGTATCCTCTCTCTTCTTGTTGGTCAACACTGAFTCAGAAAC	591
66	Db	TCTATAATCGTCTTACCANATACGTATACCTCGATTCGCATCAATCTGAATCTGAGAA	125
592	Qy	TCAACTAAATGTATCTAGAGTACAGTATGTGATCGGGTTTCATATGTACCATCGGTGCTTCC	651
126	Db	ACAAGAACCTTTCCAAACAGAAACAATAATGGGTTCTCTGTGCCTAGGTGCATCT	185
652	Qy	GCTGGGATTGGAATGCTTACTATCTCTGATACAATAGCTTTCAGAGAAAGTTTTTCACGAAG	711
186	Db	GCCATATTTGCAATGTCATCACTCTCTTATGCAAATTTTATTTTGAGAAAAATTAATAAGACT	245
712	Qy	CATACATCTCAGCAGTACAGGACTTGGCCATTTACCAAGTCTCTAGTTGCGAGTTCTGTGA	771
246	Db	GAAACTTTTTCTACTGTATTAAGCAATGATTTTTTACCCAATGATCTGTGGGTGCAATTTGGT	305
772	Qy	GTTCTCATAGGACTTTTTGCAAGTGGAGATGGGAAACTTTGCCAAGTGAGATGAGAAAC	831
306	Db	GGCCTTGTGGGATGCTTGTAAGTGGAGATTTGGAGAACTATGGGAATGGAGATGAAGGAA	365
832	Qy	TACAAACTCGGAAAGTGCATATGTTTTGACTTTTAGCTCGGACGCTATTTCCTGGCAA	891
366	Db	TTCCGAAACGGTAGTGTGCATATGTAATGACTCTGGTTTGTACTTCTGTGACATGGCAG	425
892	Qy	GTCTACACTCCTGGTCTTGTGGGATTGATCTTCAGAGTCATCCTCTGTGTTTCCCAATTC	951
426	Db	ATAGGTTGTGTTGGTATGCTGGGTTGATTTTTGAGGTATCTTCATTACCTCAGCTGC	485
952	Qy	A-TAACAGCTGTGGGATTGCCTATAGTTCC	980
486	Db	ACTTAAGCAATCTGGAATTGAACATAGCTCC	515

RESULT 11

US-10-424-599-9213

US-10-424-333-3213  
: Sequence 9213, Application US/10424599

; sequence 3213, Application US/1  
: Publication No. US20040031072A1

: GENERAL INFORMATION:

APPLICANT: La Rosa Thomas J

APPLICANT: LA ROSA THOMAS J  
APPLICANT: KOVALIC DAVID K

APPLICANT: Zhou Yihua

APPLICANT: ZHOU YINHA  
APPLICANT: CAO YONGWANG

APPLICANT: Cao Yongwei

TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With

1. TITLE OF INVENTION: PLANTS AND USES THEREOF FOR PLANT IMPROVEMENT

1. TITLE OF INVENTION: Plants and Uses Thereof

FILE REFERENCE: 38-21(53223)B

; CURRENT APPLICATION NUMBER: 08-06977

; CURRENT FILING DATE: 01/01/2010

; NUMBER OF S

; SEQ ID NO 9213

; LENGTH:

; TYPE: DNA

; ORGANISM: Glycine ma

; FEATURE:

NAME/KEY	VALUE
NAME	NAME
KEY	KEY

; LOCATION: (1)..(792)

OTHER INFORMATION:

Query Match	12.9%	Score 154.4;	DB 16;	Length 792;
Best Local Similarity	54.9%;	Pred. No. 8.4e-37;		
Matches 333;	Conservative	0;	Mismatches 262;	Indels 12; Gaps 1;

Qy	169	TGCCAGCCACTAGCTACAAATCTTGGGTAGATGTGTAATAAGAAATGGAGGAAATAGACACA	228
Db	161	TGCCAATGTTTCAGGCCACTCTCTTTTGGGAAGGTTTTTACTTTTGACAAAGGTGGTAAAAAGCAAA	220
Qy	229	TATGTGTTAGACACTTCTTTCAACTCAATTTGGCTTCCCTGTACTGGTTCCTGTCGGCTTCCTT	288
Db	221	TGGATTTGAAGCAATTTGTTTCAATTCAGCTGGATTCCTCTATACAAATTCCACTCCTCTATTTTAC	280
Qy	289	TCTCGAATCAGGCAACCCAAATCAACAGATACAAATTTTCAGTCAGTCC-----	336
Db	281	TGTCCTCAAAATGCAAAATCAACAAACGATCCTAGTAATAATGATACCTCCGAAACCAAA	340
Qy	337	CTTTCCTTACCACCCCTTGGCATCGGTTTACTTGTGCACTGGACTGTAGTGTCCGCTTAT	396
Db	341	CCAAAACCTTCTCATATATTTCTCTGTGACCTAGTTTGTGATTATTTGTTAGCAGCCATA	400
Qy	397	GCCTTATTTGCTGCAGTAGGGTGTCTACTTACTTACCAGTCTCTACTTCTCCCTCATCTTG	456
Db	401	GACTTTATGATGATGATGATGAGACTTTTATACCTTCTCTTCCACTTTTGCACTATTATGT	460
Qy	457	GCCTCACAGTTGGCTTTCACCTGCTCTTTTCTCATATTTCTTAACTCGCAAAAGTTCCT	516
Db	461	GCATCCCAATTTGGCTTCATTTGCACTGTGACTTTCTTTCATCAATTCCTCAAAAGTTCCT	520
Qy	517	CTTTGATAGTACGTTCTTCTGCTTCTCTCTACTGATGCTCTGCTCTCTCTTGTGGTCAAC	576
Db	521	GCATCATATTAACCTCTGTAGCTGTCTCTTAGCATGTGAGTATCTCTAAITTCGACTCAAC	580
Qy	577	ACTGATTCAGAAACTCAACTAATGATATCTAGAGTACAGTATGATCGGGTTCATATGT	636
Db	581	CTGAAATCTGAGGACACAAAGCACCATCCAGAGAGAGCAAAATAATTTGGGATCCTCTGT	640
Qy	637	ACCATCGGTGCTCCGCTGGGATTTGAGCTGTTACTATCTCTGATACAAATGCTCTTCAGG	696
Db	641	GCCATAGTCTCATCTGCGCAAGTTTGGCACTCCATCACTCTCTGTGTCAGCGTGTGTTTGAG	700
Qy	697	AAAGTTT 703	
Db	701	AAAGTAT 707	

RESULT 12

US-10-425-114-12485

; Sequence 12485, Application US/10425114

; Publication No. US2004003488A1

; GENERAL INFORMATION:

; APPLICANT: Liu, Jingdong

; APPLICANT: Zhou, Yihua

; APPLICANT: Kovalic, David K.

; APPLICANT: Screen, Steven E

; APPLICANT: Tabaska, Jack E

; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53313)B

; CURRENT APPLICATION NUMBER: US/10/425,114

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 73128

; SEQ ID NO 12485

; LENGTH: 1386

; TYPE: DNA

; ORGANISM: Zea mays

; FEATURE:

; OTHER INFORMATION: Clone ID: 701179255\_FLI

US-10-425-114-12485

Qy	475	ACTGCCTTTTCTCATATTTCTCTTAACTCGGAAGTTTCACTTCCTTTGATAGTCAGTTCT	534
Db	524	AAGGCATCACTTCGGGCTTATCAAGCGCAGCGGTTCA CGGCGTGATGCCAACATCC	583
Qy	535	TTGCTTTCCTCACGTATCCTCTCTCTTTGTGGTCAACA CATGATTC---AGAAAC	591
Db	584	GTGGTCGTGCTCACCTTCTCGCGCGATCCTTCGGCGTCGGATCTCTCTCCGACGAGACC	643
Qy	592	TCAACTAATGATCTAGAGTACAGTATGTATGGGTTCAATGTAACATCGGTGCTTCC	651
Db	644	TCCAGCAACGTGCGCGCGCAAGTACCACTGGGGTTCGTA CTGTCGTGCCGCCCTCG	703
Qy	652	GCTGGGATTTGACTGTTACTATCTCTGATACAAATGCTCTTCAGGAAAGTTTTCACGAAG	711
Db	704	CGGGTTTCGCGCTCATTTTGTCCCTCTTCGAGCTCACCTTCGAGAAGTGTCTAGSGTG	763
Qy	712	CATACATCTCAGCAGTCACGGACTTGGCCATTTACCAAGTCTCTAGTTGCGAGTTGTGTA	771
Db	764	CGGACGGCGGTCGTGGGTCCTCGGATGCAGATGCACCAACCTCGTGGTGGCTGCGCGGTG	823
Qy	772	GTTCTCATAGGACTTTTTGCAAGTCGAGAGTGGGAAAATTTCGCAAGTGAGATGAGAAAC	831
Db	824	TCCGTGGCGGGCTGCTCGCGTCGGGGACTGGCGGACGATCCCGGGGGAGATGGCGTCG	883
Qy	832	TACAAACTCGGGAAGTGTCATATGTTTGTACTTTAGCCTCGGACGTATTTCTTGGCAA	891
Db	884	TTCAAGAAGCGGAGACGAGGTACGTGTGACGCTGTGGGCGCGCGGTGTCGTGGCAG	943
Qy	892	GTCTACACTCTGTGCTTTGTGGGATTTGATCTTCGAGTCATCTCTGTGTCTCCAAATCC	951
Db	944	GCAGCGCGTGGCTTGGTCGGCTGATCATGAGGTGTCTGCTGTTTCGGAACGTG	1003
Qy	952	ATAACAGCTGTGGGATTCGCTATAGTTCAGTTCGGCGAGTGATAGTTTTCATGATAGA	1011
Db	1004	ACGTGCACGCTGGCGCTGCCGCTGGTGCAGTGTTCGGGTGGCGCTGTTTCGGGACAGG	1063
Qy	1012	ATGACGCAATCCAAATCTCTCCATTATTTTAGCTATCTCGGCTCTCTTTCATTCCTC	1071
Db	1064	ATGACCGGTATAAGATCGTGGCCATGCTCATGGCGGTATGGGGGTTTCTCTCTACATG	1123
Qy	1072	TATCAGCACTACCTCGACGAA	1092
Db	1124	TACCAGCACTACATCGACGCA	1144

RESULT 13  
US-10-425-114-28473  
; Sequence 28473, Application US/10425114  
; Publication No. US20040034888A1

```

: GENERAL INFORMATION:
: APPLICANT: Liu, Jingdong
: APPLICANT: Zhou, Yihua
: APPLICANT: Kovalic, David K.
: APPLICANT: Screen, Steven E
: APPLICANT: Tabaska, Jack E
: APPLICANT: Cao, Yongwei
: TITLE OF INVENTION: Nucleic A
: TITLE OF INVENTION: Plants a
: FILE REFERENCE: 38-21(53313)B
: CURRENT APPLICATION NUMBER: 1
: CURRENT FILING DATE: 2003-04
: NUMBER OF SEQ ID NOS: 73128
: SEQ ID NO 28473
: LENGTH: 1779
: TYPE: DNA
: ORGANISM: Zea mays
: FEATURE:
: OTHER INFORMATION: Clone ID:
US-10-425-114-28473

```

Query Match	12.1%;	Score 145;	DB 16;	Length 1779;
Best Local Similarity	51.8%;	Pred. No. 1.1e-33;		

Matches	353;	Conservative	0;	Mismatches	325;	Indels	3;	Gaps	1;
QY	415	GGGTTGCTCTACTTACAGTCTCTACTTTCTCCCTCACTCTGGCCTCACAGTTGGCCCTC	474						
Db	828	CGCGTGAGTACCTGCGCGTGTCCACCTTCTCGTCTGGCGCGACGCGAGCTGGGCTTC	887						
QY	475	ACTGCCCTTTTCTCATATTTCTTAACTCGAAAAAGTTCACTCCTTTGATAGTCAGTTCT	534						
Db	888	AACGCCATCACTCGCCCTCATCAACGCGCAGCGGTTCAACGCGCTGATCGCAACTCC	947						
QY	535	TTGCTTCTCTCACTGTATCTCTGCTCTTCTTGTGTCAACACTGATTC--AGAAAAAC	591						
Db	948	GTGGTGTGTCTACCTTCTCGCGCGCATCTCGCGTCTGGATCTCTCCGACGAGACCC	1007						
QY	592	TCAACTAATGTATCTAGAGTACAGTATGTGATCGGGTTCAATGTACCATCGGTGCTTCC	651						
Db	1008	TCCAGCAACGTGCGCGCGCAAGTACCCGCTGGGGTTCTGCTCGTCTGGCGCCTCG	1067						
QY	652	GCTGGGATTTGAGCTGTACTATCTGTATACAAATGCTCTTCAGGAAAGTTTTTCACGAAG	711						
Db	1068	CGCGTGTTCGCGCTCATTTTGTCCCTCTTCGAGCTCACCTTCGAGAAAGTGGTCAGGGTG	1127						
QY	712	CATACATCTCCAGCAGTACAGCACTTGGCCATTTACCACTCTCTAGTTGCGAGTTGTGTA	771						
Db	1128	CGGACGCGCGGTGGTGGTGGCGATGCATGCACCAACCTGTGTGGCGTGGCGGGTG	1187						
QY	772	GTTCTCATAGACTTTTTGCAAGTGGAGAGTGGGAACTTTTGCAAGTGAGATGAGAAAC	831						
Db	1188	TCCGTGGCGGGCTGCTCGCGTGGGGGACTGCGGACGATCCGGGGGAGATGGCGTGC	1247						
QY	832	TACAACTCGGGAAAGTGTATATGTTTTCAGCTTTTAGCCTTCGGCAGCTATTTCTCGCAA	891						
Db	1248	TTCAAAGACGGGAGGACAGGTACGTCTCAGCTGTGTGGGCACGGCGGTGTCTGTGCAG	1307						
QY	892	GTCTACACTCCTGGTCTTGTGGGATTTGATCTTCGAGTCATCTCTGTGTTCTTCAATTC	951						
Db	1308	GCACGCGCGTGGGCTTGGTGGCGCTGATCATGAGGGTGTCTGCTGTTTCGCGAACGTG	1367						
QY	952	ATAACAGCTGTGGATGGCTATAGTTTCCAGTTTGGCGAGTGTATGTTTTCCATGATAGA	1011						
Db	1368	ACGTGCACGCTGGGCTGCCGCTGGTCCAGTGTTTGCGGTGGCGCTGTTCGGGGACAGG	1427						
QY	1012	ATGACGCACTCCAAAATCTTCTCCATTTATTTTAGCTATCTCGGCTTCTTTTCATTGTC	1071						
Db	1428	ATGACGGTATAAAGATCGTGGCCATGCTCATGCGGTACGGGGTTCCTCTCTCTCATG	1487						
QY	1072	TATCAGCACTACCTTCGACGAA	1092						
Db	1488	TACCAGCACTACATCCACGCA	1508						

```

RESULT 14
US-10-767
; Sequence
; Publication
; General
; Application
; Application
; Application
; Title
; Title
; File Name
; Current
; Current
; Number
; Seq ID
; Length
; Type
; Organ
; Feature
; Other
US-10-767

```

```

: APPLICANT: Cabot, Tongwei
: TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
: TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
: FILE REFERENCE: 38-21(53313)B
: CURRENT APPLICATION NUMBER: US/10/425,114
: CURRENT FILING DATE: 2003-04-28
: NUMBER OF SEQ ID NOS: 73128
: SEQ ID NO 28473
: LENGTH: 1779
: TYPE: DNA
: ORGANISM: Zea mays
: FEATURE:
: OTHER INFORMATION: Clone ID: LIB4757-034-H3_FLI
US-10-425-114-28473

```

US-10-767-701-9000  
; OTHER INFORMATION: C10ne ID: SORBI-28MAY03-CLUS69833

Query Match 11.9%; Score 141.8; DB 17; Length 728;  
Best Local Similarity 55.0%; Pred. No. 6.e-33;  
Matches 278; Conservative 0; Mismatches 227; Indels 0; Gaps 0;

QY 643 GGTCTTCGCTGGGATTTGACTGTTACTATCTCTGATACAAATGCTCTTCAGGAAAGTT 702  
DB 1 GGAGCCTCAGCCACATCTCACTAATCTCTCTAATGCAAGTCACATTTGAGAAGTT 60

QY 703 TTCAGAAAGCATACATCTCTAGCAGTCACGACTTGGCCATTTACGACTCTCTAGTTGCG 762  
DB 61 ATTAAGAAAGGAAACGTTCTAGCTGTTGTTGAATATGCAGATATATACAGCACAATAGTGSCA 120

QY 763 AGTTGTGTAGTTCTCATAGGACTTTTTCGAAGTGAGAGTCGGAACCTTTGCCAAGTGAG 822  
DB 121 ACAATAGCTTCTCTGTGTTGGTTATTTGCAAGCGTGAGTGAAGACTTTAGAGGGAGAG 180

QY 823 ATGAGAAACTACAACTCGGAAAGTGTCATATGTTTGGACTTTAGCCTCGGCAGCTATT 882  
DB 181 ATGCATGCTTCAGCTCAGGGAGGTTGTATATATGTATGATGACACTTCTATGGACTGCTGTA 240

QY 883 TCCTGGCAAGTCATACCTCTGCTGCTTGTGGGATTTGATGATCCTCTGTGTGTTTC 942  
DB 241 TCTTGGCAGATAGCTTCTGTAGGAGTGTGGGTTTGATCTTTTGTGTCTATCACTCTTT 300

QY 943 TCCAAATCCATAACAGCTGTGGGATTTGCTATAGTTCAGATTGCCAGTTCGCGAGTCATAGTTTC 1002  
DB 301 TCAAAATGTGAAGCACACTAGCTCTACCCATCATTCAGTTTTCGCTGTGATTTTCTTC 360

QY 1003 CATGATAGATGGAGCGCATCCAAATCTCTCCATTTATTTAGCTATCTGCGGCTTCCTT 1062  
DB 361 CACGACAGATGAATGGAGTGAAGATTATGCTATGTTGATGGCACTCTGGGATTTGTT 420

QY 1063 TCATTCTGCTATCAGCACTACTCTGCAAGAAAGATTGAATATCTAGCCACACAAAGTCT 1122  
DB 421 TCATATGATACCAATTATATATTAGTGACAAGAAAGCTAGGAAGACTTCAGTAAGTGTG 480

QY 1123 GTAGGAGATCTTCATCTACCTGTTG 1147  
DB 481 GAGGAGAATTCCTAAGTCTGTTG 505

RESULT 15  
US-10-767-701-6806  
; Sequence 6806, Application US/10767701  
; Publication No. US20040172684A1  
; GENERAL INFORMATION:  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement  
; FILE REFERENCE: 38-21(53535)B  
; CURRENT APPLICATION NUMBER: US/10/767,701  
; CURRENT FILING DATE: 2004-01-29  
; NUMBER OF SEQ ID NOS: 63128  
; SEQ ID NO 6806  
; LENGTH: 747  
; TYPE: DNA  
; ORGANISM: Sorghum bicolor  
; FEATURE:  
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS55663\_1  
US-10-767-701-6806

Query Match 11.6%; Score 138.8; DB 17; Length 747;  
Best Local Similarity 56.3%; Pred. No. 5.7e-32;  
Matches 260; Conservative 0; Mismatches 202; Indels 0; Gaps 0;

QY 107 CTCAAACCGAGCACTATAAAGGTGGCTTCGTGCTCCATATACGTAATCTTTGCTCTCT 166  
DB 285 CCCATACAGGCAATTTGGCATTTGGTGGTGGTAGCTCTGAACATTTTCTCTCTCGTTG 344

QY 167 TTTGCGAGCCACTAGCTACAATTTCTGGGTAGATTGTTACTATGAAATGGAGAAATAGCA 226

DB 345 CTGCTCAGACAGCATCGACACTCTCTGGTAGTTCTTACTACAACCAAGGTGGAACAGCA 404

QY 227 CATATGTGGTAACACTTCTTCAACTCATTTGGCTTCCCTGTACTGGTTCTGTTCGGCTTCT 286

DB 405 AATGGATGTCCACATTTTGTCCAAACCGCTGGCTTTCAGTATTTGTTGTGCGCCCTATATC 464

QY 287 TTTCTCGAATCAGGCAACCCAAATCAACAGATACAAATTTTCAGTCAGTCCCTTCCCTTCA 346

DB 465 TGTTCCTTCAAAGTCGCCCTTCTACTCAAACAACCAAGTAACCCCTGAGACTTCTGTCA 524

QY 347 CCACCCCTTGCATCGGTTTACTTTGTGCACTGGACTGTAGTGTCCGCTTATGCTTTATTTGT 406

DB 525 CCAGAATTACTCTTATATACATTTGGCTTGGGCCCTCATCATTTGCTGCCGATGACTTGATGT 584

QY 407 CTGCAAGTAGGTTGCTCTACTTACCAGTCTCTACTTTCTCCCTCATCTTGGCCTCAGT 466

DB 585 ATTCCTATGGCCTCTTGTACTCTTCCCTGTATCAACATATTCGCTCATATGCGCTAGTCAGC 644

QY 467 TGGCCTTCACTGCTCTTTTCTCATATTTCTTAACCTGCAAAAGTTTCACCTCTTGTATAG 526

DB 645 TGGCCTTCAATGCTGCTTCTCTTATGCTCTGATGCTCAAAAGTTTCACCTGATGATCC 704

QY 527 TCAGTTCTTTGCTTCTCTCTCACTGTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 568

DB 705 TCAACTCGGTAATCTCTCTTACCTTTTCTGCTGCACTTCTTG 746

Search completed: November 1, 2004, 23:40:06  
Job time : 622.418 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 1, 2004, 05:48:03 ; Search time 110.151 Seconds  
(without alignments)  
7704.726 Million cell updates/sec

Title: US-09-913-767-5  
Perfect score: 1194  
Sequence: 1 tcatgatataataaacat.....tgtgatcaagcatatttcc 1194

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA.\*  
1: /cgn2\_6/prodata/1/ina/5A COMB.seq.\*  
2: /cgn2\_6/prodata/1/ina/5B COMB.seq.\*  
3: /cgn2\_6/prodata/1/ina/6A COMB.seq.\*  
4: /cgn2\_6/prodata/1/ina/6B COMB.seq.\*  
5: /cgn2\_6/prodata/1/ina/PCBUS COMB.seq.\*  
6: /cgn2\_6/prodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	80.4	6.7	7218	1	US-08-232-463-14
2	45	3.8	1141	4	US-09-806-708B-22
3	41.6	3.5	289	3	US-09-007-005-17
4	41.6	3.5	289	3	US-09-244-796-17
5	37	3.1	399	4	US-09-621-976-8976
6	36.8	3.1	832	4	US-09-621-976-2813
7	35.8	3.0	1141	4	US-09-806-708B-22
8	35.2	2.9	248	3	US-09-007-005-32
9	35.2	2.9	248	3	US-09-244-796-32
10	35.2	2.9	277	3	US-09-007-005-3
11	35.2	2.9	277	3	US-09-244-796-3
12	34.8	2.9	1923	4	US-09-134-000C-466
13	34.4	2.9	832	4	US-09-621-976-2813
14	34.4	2.9	1318	4	US-09-719-108-1
15	34.4	2.9	1830121	4	US-09-557-884-1
16	34.4	2.9	1830121	4	US-09-643-990A-1
17	34.4	2.9	1830121	4	US-10-329-960-1
18	34.2	2.9	474	4	US-09-621-976-18033
19	34.2	2.9	591	4	US-09-495-050A-104
20	34.2	2.9	2412	4	US-09-620-312D-1023
21	34.2	2.9	4211	3	US-09-004-838-106
22	33.8	2.8	135	4	US-09-513-999C-33549
23	33.8	2.8	1860	4	US-09-919-497-50
24	33.6	2.8	282	3	US-09-461-697-205
25	33.6	2.8	306	3	US-09-461-697-203
26	33.6	2.8	696	3	US-09-461-697-193
27	33.6	2.8	699	3	US-09-461-697-191

C 28	33.6	2.8	717	3	US-09-461-697-189	Sequence 189, App
C 29	33.6	2.8	774	3	US-09-461-697-187	Sequence 187, App
C 30	33.6	2.8	801	3	US-09-459-956-7	Sequence 7, Appli
C 31	33.6	2.8	819	3	US-09-461-697-185	Sequence 185, App
C 32	33.6	2.8	1569	3	US-09-461-697-184	Sequence 184, App
C 33	33.6	2.8	4026	4	US-09-248-796A-4971	Sequence 4971, Ap
C 34	33.2	2.8	1830121	4	US-09-557-884-1	Sequence 1, Appli
C 35	33.2	2.8	1830121	4	US-09-643-990A-1	Sequence 1, Appli
C 36	33.2	2.8	1830121	4	US-10-329-960-1	Sequence 1, Appli
C 37	32.8	2.7	1298	3	US-08-948-705-3	Sequence 3, Appli
C 38	32.8	2.7	1298	4	US-09-510-543-3	Sequence 3, Appli
C 39	32.6	2.7	245	4	US-09-513-999C-10942	Sequence 10942, A
C 40	32.6	2.7	530	4	US-09-270-767-30636	Sequence 30636, A
C 41	32.6	2.7	911	1	US-08-393-985-5	Sequence 5, Appli
C 42	32.6	2.7	929	4	US-09-270-767-14462	Sequence 14462, A
C 43	32.4	2.7	685	4	US-09-620-312D-806	Sequence 806, App
C 44	32.4	2.7	3635	2	US-08-553-436A-5	Sequence 5, Appli
C 45	32.2	2.7	553	4	US-09-621-976-15491	Sequence 15491, A

ALIGNMENTS

RESULT 1  
US-08-232-463-14  
; Sequence 14, Application US/08232463  
; Patent No. 5670367  
; GENERAL INFORMATION:  
; APPLICANT: DORNER, F.  
; APPLICANT: SCHEIFLINGER, F.  
; APPLICANT: FALKNER, F. G.  
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 1800 Diagonal Road, Suite 500  
; CITY: Alexandria  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22313-0299  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/232,463  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/935,313  
; FILING DATE:  
; APPLICATION NUMBER: EP 91 114 300.6  
; FILING DATE: 26-AUG-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BENT, Stephen A.  
; REGISTRATION NUMBER: 29,768  
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703)836-9300  
; TELEFAX: (703)683-4109  
; TELEX: 899149  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7218 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; CLONE: pTZgpt-F1s  
US-08-232-463-14

Query Match 6.7%; Score 80.4; DB 1; Length 7218;



Db 165 YSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYN 106  
QY 452 TCTGGCCTCACAGTTGGCCTTCACGCGCTTTTCTCATATTTCTTAACTCGCAAAAGT 511  
Db 105 YSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYN 46  
QY 512 TCACTCCTTTGATAGTCAGTTCTTTGCTCTCTCTCACT 549  
Db 45 YAYAYTTYGYTAYAYAYTAYGYTAYAYTTYGYT 8

RESULT 4  
US-09-244-796-17/c  
; Sequence 17, Application US/09244796  
; Patent No. 6281344  
; GENERAL INFORMATION:  
; APPLICANT: Szoatak, Jack W.  
; APPLICANT: Roberts, Richard W.  
; APPLICANT: Liu, Rihe  
; TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN  
; FILE REFERENCE: 00786/350007  
; CURRENT APPLICATION NUMBER: US/09/244,796  
; CURRENT FILING DATE: 1999-02-05  
; EARLIER APPLICATION NUMBER: 60/035,963  
; EARLIER FILING DATE: 1997-01-27  
; EARLIER APPLICATION NUMBER: 60/064,491  
; EARLIER FILING DATE: 1997-11-06  
; EARLIER APPLICATION NUMBER: 09/007,005  
; EARLIER FILING DATE: 1998-01-14  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 17  
; LENGTH: 289  
; TYPE: RNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Translation template  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)-(289)  
; OTHER INFORMATION: n = A,T,C or G

US-09-244-796-17  
Query Match 3.5%; Score 41.6; DB 3; Length 289;  
Best Local Similarity 10.4%; Pred. No. 0.004;  
Matches 29; Conservative 108; Mismatches 141; Indels 0; Gaps 0;  
QY 272 TTCTGTTCGCTTCTTTTCTCGAATCAGGCAACCCAAATCAACAGATACAAATTCAGTC 331  
Db 285 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 226  
QY 332 AGTCCCTTCTTCACACCCCTTGATCGTTTACTTTGTCAGTGCAGCTAGTGTCCG 391  
Db 225 YCYTGYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYN 166  
QY 392 CTATCTCTATTGTCTGAGTAGGCTGCTCTACTTACAGTCTCTCTCTCTCTCA 451  
Db 165 YSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYN 106  
QY 452 TCTGGCCTCACAGTTGGCCTTCACGCGCTTTTCTCATATTTCTTAACTCGCAAAAGT 511  
Db 105 YSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYN 46  
QY 512 TCACTCCTTTGATAGTCAGTTCTTTGCTCTCTCTCACT 549  
Db 45 YAYAYTTYGYTAYAYAYTAYGYTAYAYTTYGYT 8

RESULT 5  
US-09-621-976-8976/c  
; Sequence 8976, Application US/09621976  
; Patent No. 6639063

; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Jobert, S.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
; FILE REFERENCE: GENSET.054PR2  
; CURRENT APPLICATION NUMBER: US/09/621,976  
; CURRENT FILING DATE: 2000-07-21  
; NUMBER OF SEQ ID NOS: 19335  
; SOFTWARE: Patent.pm  
; SEQ ID NO 8976  
; LENGTH: 399  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-621-976-8976

Query Match 3.1%; Score 37; DB 4; Length 399;  
Best Local Similarity 12.5%; Pred. No. 0.14;  
Matches 46; Conservative 162; Mismatches 157; Indels 2; Gaps 1;  
QY 684 AATGCTCTTCAGAAAGTTTCCAGAAAGATACATCCTCAGCAGTCACGCGACTTGGCCAT 743  
Db 365 AATCAGCAAGKWWGAMKWTGKWKYVYVYVYVYVYVYVYVYVYVYVYVYV 306  
QY 744 TTACCACTCTCTAGTTGCGAGTTGTAGTTCTCATAGACTTTTTCGCAAGTGGAGAGTG 803  
Db 305 KXWYV 246  
QY 804 GGAACCTTTCAGAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGT 863  
Db 245 SWYSAMRRKKKWTCKWGRS--SWGSRSTGYAYAMWKSWCTSRKWKYVYVYVYVYVYVYV 188  
QY 864 TTTAGCTCGGCGAGCTATTTCTGCGCAAGTCTACACTCTGCTCTGCGGATTCATCTT 923  
Db 187 STKTCYRGSTYCKWKAYTKRKRKWTWYVYVYVYVYVYVYVYVYVYVYVYVYV 128  
QY 924 CGAGTCATCTCTGTTCTTCCATTAACAGCTGTGGAGTTCCTCATATAGTTCACAT 983  
Db 127 TKWTCTWCKCTTYMAGTMMYV 68  
QY 984 TCGGCGAGTGCAGTTCCTCATGATAGATGAGCAGCAGTCCAAATCTTCTCCATTTT 1043  
Db 67 WMSMSMKWKGSMWKKWTYYYYYMMKWSKMTYWSMNSCYARKKWRVYAKTYTWTMTW 8  
QY 1044 AGCTATC 1050  
Db 7 GKRTYKY 1

RESULT 6  
US-09-621-976-2813  
; Sequence 2813, Application US/09621976  
; Patent No. 6639063  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Jobert, S.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
; FILE REFERENCE: GENSET.054PR2  
; CURRENT APPLICATION NUMBER: US/09/621,976  
; CURRENT FILING DATE: 2000-07-21  
; NUMBER OF SEQ ID NOS: 19335  
; SOFTWARE: Patent.pm  
; SEQ ID NO 2813  
; LENGTH: 832  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 235..399  
US-09-621-976-2813

Query Match 3.1%; Score 36.8; DB 4; Length 832;

```
Best Local Similarity 11.8%; Pred. No. 0.24;
Matches 41; Conservative 149; Mismatches 156; Indels 0; Gaps 0;

QY 830 ACTACAACTCGGGAAGTGCATATGTTTGGACTTTAGCCTCGGCACTATTTCTCGGC 889
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 26 WMYKWKYMYKTYWRWRKKKAWKWKYKWTWYRYVAMWGTYKKKAMCRKTKKKKKG 85

QY 890 AAGTCTACACTCCCTGGTCTTCTGGGATGTATCTTCGAGTCATCTCTGTCTTCCAATT 949
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 86 YMMWYMGWRSSYAMWMTRTWTGAYYRSMYMYWRYRCWKKKAYRKTCTSYSSKGWTKW 145

QY 950 CCATAACAGCTGGGATTGCTATAGTTCAGTTGGCGAGTATAGTTTTCATGATA 1009
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 146 RWKAWTWWKKYKYWAATYVMMWCKWKWRASWYCWGWGKARKWSWTKRSRVSASA 205

QY 1010 GAATGGAGCATCAAAATCTCTCCATATATTTAGCTATCTGCGGCTCTCTTCATCG 1069
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 206 RSAKRCYSCSWGAMSWKMYMRWRWGWATGAGKAWRASCMWRKYAGKSKTSYKSMW 265

QY 1070 TCTATCAGCACTACTCGACGAAAAGAGTTGAATACTAGCCACACAAGTCTGTAGAG 1129
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 266 MCWTRSMKYCTKARWTGYCYRKGMMGKGRWYASKYKMYKSWMMWCMWARMYRYSTGR 325

QY 1130 ATCTTCACTACCTGTTGAGGAAGTGCACAAACATACAAAGTGT 1175
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 326 ASMWRRRYHYTMKKWKYAWARAARWAWWAWWRACAAATAAT 371

RESULT 7
US-09-806-708B-22
; Sequence 22, Application US/09806708B
; Patent No. 6784342
; GENERAL INFORMATION:
; APPLICANT: The University of British Columbia
; TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
; FILE REFERENCE: 4810-58741
; CURRENT APPLICATION NUMBER: US/09/806,708B
; CURRENT FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/147,133
; PRIOR FILING DATE: 1999-08-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22
; LENGTH: 1141
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (1)..(1141)
; OTHER INFORMATION: consensus sequence of A.t., L.a., and B.n. PAF1 promoters
US-09-806-708B-22

Query Match 3.0%; Score 35.8; DB 4; Length 1141;
Best Local Similarity 11.5%; Pred. No. 0.6;
Matches 84; Conservative 242; Mismatches 395; Indels 7; Gaps 1;

QY 399 TTATTTGTCGAGTAGGGTGCTCTACTTACCAAGTCTCTACTTCTCCCTCATCTGGC 458
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 309 WKSRTTKVRTSCWANNCRAGDANKHKWKKWSAAMVYNNNNNNNTYKKARHBRWDW 368

QY 459 CTCACAGTTGGCCTTCACTGCTTTTCTCATATTTCTTAACTCGCAAAAGTTCACCTC 518
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 369 VWHSAWKWANAHAHYSRKKWTBTKRTWNNNGTTWKKRWAWYWKMDMDWBGTTNNN 428

QY 519 TTTGATAGTCAGTTCTTCTCTCCTCATGTATCTCTCTCTCTCTCTCTCTCTCTCTCT 578
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 429 NNGGRTYGVTKNKKMWYTKWANNCKWRADWHTCTHNNTTWWMKTYWNNCYWKSMT 488

QY 579 TGATTCAGAAAACCAACTAATATGTATCTAGAGTACAGTATGTATCGGGTTCATATGTAC 638
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 489 NGKSHRBAAYVYTMWMMRRYAHANNNNNDWYWKACTWYKVCSSKWNWNNYAAWTKSS 548

QY 639 CATCGGTCTTCCGCTGGGATGGACTGTTACTTATCTCTGATACAAATGCTCTTCAGGAA 698
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
549 WNTSRYRWKTNNSWRSDTRSMGRANNYARABHYGKWNTRWBSHTWBBRAGAA 608
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
699 AGTTTTCAGAAACATACATCTTCAGCAGCTCAGGACTTGGCCATTTACCAAGTCTTAGT 758
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
609 HYWBMWYBAKCHCKAWYKAKKYAGAGSNNNNNNNNNNNNNNNNNNNNATCADDYYAASRW 668
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
759 TGGAGTGTGTAGTCTCATAGGACTTTTTCGAAGTGGAGAGTGGGAACCTTTGCCAAG 818
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
669 YAMANAKWYKYBAANNAYYTHANNWGCWNNATDTRRTWKKNNNNNNAGTWKNNNNNA 728
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
819 TGAGATGAGAAATTCACAACTCGGAAAGTGTCTATATGTTTTCG-----CTTTAGCCT 871
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
729 KNSAAKHYAAAAYAAKAKKHWKANKWAGWGHADAAABTTDKNNGAYTKYTTNNNT 788
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
872 CGGAGCATTTTCTCGCAAGTCTACACTCTCTGCTGTGGGATGTATCTTCGAGTCAT 931
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
789 YRGVVTNTAARDGWANNNNNNNNNNNNNNNNNNNNNGSDMWVTWYAYNYGTTNNNNNNNNAYAW 848
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
932 CCTCTGTGTTCTCCAATTCATAACAGCTGTGGGATTCCTATAGTTCAGTTCCAGTTGCGGCAG 991
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
849 WTKWYTTDDRWBRBATYNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNAT 908
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
992 TGATAGTTTCCATGATAGAAATGAGCGCATCCAAAATCTTCTCCATTTATTTAGCTATCT 1051
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
909 YNRGTARNTNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNCT 968
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1052 GCGGCTTCTTCAATTCATGCTATCAGCACTACTCGACGAAAAGAGTTCGAATAGTACGC 1111
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
969 TTCRKYNKNCWTYTMWTTTTRTWYAAATRWKTNNTATGSMTRCNATGKNNNYWTGKTRWT 1028
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1112 ACACAGT 1119
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1029 AYRMATRW 1036
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 8
US-09-007-005-32/c
; Sequence 32, Application US/09007005B
; Patent No. 6258558
; GENERAL INFORMATION:
; APPLICANT: Szostak, Jack W.
; APPLICANT: Roberts, Richard W.
; APPLICANT: Liu, Bihe
; TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN
; TITLE OF INVENTION: FUSIONS
; FILE REFERENCE: 00786/350003
; CURRENT APPLICATION NUMBER: US/09/007,005B
; CURRENT FILING DATE: 1998-01-14
; EARLIER APPLICATION NUMBER: 60/035,963
; EARLIER FILING DATE: 1997-01-27
; EARLIER APPLICATION NUMBER: 60/064,491
; EARLIER FILING DATE: 1997-11-06
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 32
; LENGTH: 248
; TYPE: RNA
; ORGANISM: Homo sapiens
US-09-007-005-32

Query Match 2.9%; Score 35.2; DB 3; Length 248;
Best Local Similarity 19.2%; Pred. No. 0.38;
Matches 40; Conservative 80; Mismatches 88; Indels 0; Gaps 0;

QY 241 CTCTTCAACTCATTTGCTTCCCTGCTACTGTTCTGTTCGCTTCTTTCTCGAATCAGG 300
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 212 CYTYGYTYTCYCAVGYTYTYGYTYTYTYTYTYTYTYTYTYTYTYTYTYTYTYTYTY 153

QY 301 CAACCCAAATCAACAGATACAAATTCAGTCAGTCCCTTCTTCCACCACCTTGCATCG 360
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 152 AYCYGYTYTYTYCAVGYCYCAVGYGYTYTYTYTYTYTYTYTYTYTYTYTYTYTYTYTY 93
```







**THIS PAGE BLANK (REV. 10/1/79)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 1, 2004, 05:45:29 ; Search time 3801.29 Seconds  
(without alignments)  
10228.408 Million cell updates/sec

Title: US-09-913-767-10  
Perfect score: 1067  
Sequence: 1 agacaagaatggtgaaggct.....gaccataagtgaagttaa 1067

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:  
1: gb\_est1:  
2: gb\_est2:  
3: gb\_hc:  
4: gb\_est3:  
5: gb\_est4:  
6: gb\_est5:  
7: gb\_est6:  
8: gb\_gsa1:  
9: gb\_gsa2:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	440	41.2	1230	CNS0A99D	BX821376 Arabidopsis
2	381	35.7	989	CL479645	SAIL 30 B
3	342.8	32.1	1059	BX841065	BX841065
C 4	333.2	31.2	827	BZ515292	BOMQ20TR
5	306	28.7	680	BG439880	GA_Ea000
C 6	293.2	27.5	688	BH589732	BOGES04TR
7	284.4	26.7	971	CK272858	EST18936
8	268.4	25.2	986	CK243751	EST727388
9	265.8	24.9	985	CK257466	EST741103
10	265.2	24.9	688	AJ807511	AJ807511
11	264.8	24.8	1013	CK251896	EST735533
12	262.8	24.6	998	CK247918	EST731555
13	261.2	24.5	907	CK257887	EST741524
14	246.6	23.1	988	CK248243	EST731880
15	246	23.1	979	CK252005	EST735842
16	242.4	22.7	722	BM436674	VVA008803
17	238.4	22.3	645	AU238506	AU238506
18	236.6	22.2	703	CB006438	VVC033611
19	236.4	22.2	930	CK253429	EST737066
20	235.6	22.1	942	CK249994	EST733631
21	234.8	22.0	929	CK249872	EST733509
C 22	234.4	22.0	868	CK245030	EST728667
C 23	234.4	22.0	902	CK262518	EST708596
24	231.6	21.7	908	CK247701	EST731338

25	227.4	21.3	932	7	CK256828	EST740465
26	227	21.3	665	6	CD714562	VVB202D06
27	225.8	21.2	668	6	CB008408	VVC051C02
C 28	225.4	21.1	527	8	AQ958362	LERAX12TR
29	222.2	20.8	928	7	CK245031	EST728668
30	222	20.8	948	7	CK257752	EST741389
31	221.2	20.7	784	8	BH601144	BOHJF33TF
32	219.8	20.6	829	7	CK262517	EST708595
33	219.2	20.5	899	7	CK256093	EST739730
34	218.6	20.5	894	7	CK247243	EST730880
C 35	217.8	20.4	665	5	BQ799268	EST 1437
36	217	20.3	882	7	CK258020	EST741657
37	217	20.3	889	7	CK251126	EST734763
38	216	20.2	880	7	CK257437	EST741074
39	215.2	20.2	616	5	BQ873555	BQ873555 QG1N19.Y
40	215.2	20.2	701	5	BQ874152	BQ874152 QG1A120.Y
41	214.2	20.1	892	7	CK246384	EST730021
42	213.8	20.0	732	5	BQ988613	BQ988613 QGF15G06.
C 43	212	19.9	628	5	BQ798564	EST 733 G
C 44	209.8	19.7	687	5	BQ798954	EST 1123
45	208.2	19.5	257	8	BH812211	SALK_0614

ALIGNMENTS

RESULT 1  
CNS0A99D  
LOCUS Arabidopsis thaliana Full-length cDNA Complete sequence from clone GSLTSL372D07 of Silique of strain col-0 of Arabidopsis thaliana (thale cress).  
ACCESSION BX821376  
VERSION BX821376.1 GI:42468600  
KEYWORDS HTC; GSLT cDNA.  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
REFERENCE 1 (bases 1 to 1230)  
AUTHORS Castellani, V., Aury, J.M., Jaillon, O., Wincker, P., Clepet, C., Menard, M., Cruaud, C., Quetier, P., Scarpelli, C., Schachter, V., Temple, G., Caboche, M., Weissenbach, J., and Salanoubat, M.  
TITLE Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences: A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 1230)  
AUTHORS Genoscope.  
TITLE Direct Submission  
JOURNAL Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)  
COMMENT The sequences are based on single pass reads.  
Life Technologies (a division of Invitrogen) members carried out full-length libraries construction : Temple G.  
Genoscope members carried out sequencing and annotation : Castelli V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M.  
URGV INRA : Clepet C., Caboche M.  
Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich Information center for Protein Sequences). 5 prime and 3 prime are assembled with Phrap.  
http://www.genoscope.cns.fr/externe/sequences/Banque\_Projet\_EF/Full length  
http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.

FEATURES  
source  
1..1230  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/strain="Col-0"  
/db\_xref="taxon:3702"







/clone_lib="Gossypium arboreum 7-10 dpa fiber library" /note="vector: pBK-CMV; Site_1: EcoRI; site_2: XhoI"									
ORIGIN	Query Match	28.7%;	Score 306;	DB 4;	Length 680;				
	Best Local Similarity	66.6%;	Pred. No. 8.4e-75;						
	Matches 454;	Conservative 0;	Mismatches 225;	Indels 3;	Gaps 1;				
QY	289	ACTACTTGTATGCATATGGTATAGCTATCTTCCAGTTCTTCACAGCTGCTCTTATCATTTG	348						
Db	1	ACTACTTCTACTCTCATGGCTTATCTCGTCTCCCGTTTCGACTTCTCTTTTGATCATCG	60						
QY	349	CTTTCACAGTTAGCTTTTATAGCTATCTTCTCATCTTCATGTTTAAACATAAAGTTCACTC	408						
Db	61	CGTCGCAAGTTGGCTTTTCACGGCGGGTTTGCTTCTGTTGGTGAACAAAAGTTTCACTG	120						
QY	409	CTTTTACCATCAATGCTGTTGTGTTGTGACTGTTTGGTGTGCTGCGGTTTGGGAATGCATA	468						
Db	121	CCTACTCCATAAAGCCCGTGTGTTTGTGACCATAGGGGCTGGCGGTTCTTGGCTTTGCATT	180						
QY	469	CCGAAACTGATAAGCCAGTTTCATGAGACTCACAAAGCAGTACATAAATCTGTTTCTTGATTA	528						
Db	181	CGAGCAGCAGCCGCTCGAATGAATCCAAAGAGATATATTTTGGGGTTCGTAAATGA	240						
QY	529	CTGTAGCAGAGCTGTGTATGTATGCTTTTCATCTTTGCCATTTAGTGGAACTTGCCTTACCAGA	588						
Db	241	CCTTAGCTGCAGCGGCTTTGTACGGATTTGTATTGGCTTTTGGTGGAACTAAACGTACAAGA	300						
QY	589	AACTAAGCAAAACCATGAGCTATACCTCTGTGCTCGAGTTCAGTTTGAATTTGTGTCTCC	648						
Db	301	AGGCAAAAGCAAGAGATCAGCTACGCCCTTGTGATGGAGATTACAGCTGGTGAATGTGTTGG	360						
QY	649	TTGCTTCTATTGTCCAGCGTCATCGGTATGTTTCATCGCTGGTGATTTCAAGCAGGCGCTTAC	708						
Db	361	TTGCTACTGCTTTTGCACGGTTGGGATGCTGTTCAACATGATTTCA--AGTGATCG	417						
QY	709	CAAAAGAACCAAGAGCTTCAAGCTTGGAGAGCATTTGTTCTATGTGGTGGCTGTGTTTT	768						
Db	418	GAAGGAAGCAAGGGAATTCGAGCTAGGAGAAAACAAATACTATGTGGTTGTGATTTTGA	477						
QY	769	CAGCATCATATGGCAAGGCTTCTTCTTGGAGCCATTCGATTAATCTTCTCCACATCGT	828						
Db	478	CTGCAATAATAAATCAATGTTTCTTCTTGGAGCTATTGGAGTAGTATTTTGTGCGTCAT	537						
QY	829	CTCTCGTCTCGGTATTATGATATCATGCTGCTTTTGCCAAATPACAGAGGTTTTAGCTGTGA	888						
Db	538	CAATGCTATCGGTGTAGTGATAGCGGTTCTATTACCAGTTTACAGAGATTTTGGCAATAT	597						
QY	889	TATTCTACCAATGAAAAGTTTCAAGCTGAGAAGGGACTTTCTCTTGCTCTCTCCCTTTGGG	948						
Db	598	TTTCCCTTCAATGAGAAGTTTCAAGCCGAGAAAAGGTGTCTCTCTTGTCTCTCGCTGTGA	657						
QY	949	GCCTTGTCTTACTTTTATGG	970						
Db	658	GCTTCGATAACTACTTTTATGG	679						
RESULT 6									
BH589732/c									
LOCUS	BH589732	688 bp	DNA	linear	GSS 15-DEC-2001				
DEFINITION	BOGES04TR BOGE Brassica oleracea genomic clone BOGES04, genomic survey sequence.								
ACCESSION	BH589732								
VERSION	BH589732.1 GI:17842184								
KEYWORDS	GSS.								
SOURCE	Brassica oleracea								
ORGANISM	Brassica oleracea								
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;									
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;									
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.									
1 (bases 1 to 688)									
REFERENCE	Town, C.D., Van Aken, S., Utterback, T., Koo, H. and Fraser, C.M.								
AUTHORS	Whole genome shotgun sequencing of Brassica oleracea								
TITLE									

JOURNAL COMMENT

Unpublished (2001)  
Other GSSs: BOGES04TF  
Contact: Chris Town

TIGR  
9712 Medical Center Drive, Rockville, MD 20850, USA.  
Tel: 301-838-3523  
Fax: 301-838-0208  
Email: cdtown@tigr.org  
DNA is from a doubled haploid provided by Tom Osborn.  
Seq primer: TR  
Class: sheared ends.

FEATURES  
source

1. .688  
/organism="Brassica oleracea"  
/mol\_type="genomic DNA"  
/strain="TO1000DH3"  
/db\_xref="taxon:3712"  
/clone="BOGES04"  
/clone\_lib="BOGE"  
/note="Vector: pHOS1; Site 1: BstXI; 2-3 kb sheared genomic DNA inserted into pHOS1 using BstXI linkers"

ORIGIN

Query Match 27.5%; Score 293.2; DB 8; Length 688;  
Best Local Similarity 66.4%; Pred. No. 3.5e-71;  
Matches 437; Conservative 0; Mismatches 218; Indels 3; Gaps 1;

QY 2 GACAGAATGGTGAAGGCTCTTGTGATCATTAACCTGCAATAATCTAGCCATAGGAACCTG 61  
Db 662 GATGAAGATGAAGCGGTTCTTGTAAATCAATAAATGATATATATGGCGATAGGAACCTG 603  
QY 62 TGGAGGCTCTTTGATTATGCGTCTCTACTTCAACAATGGCGTAAAGGATTTGGTTCTC 121  
Db 602 TGAGGGCCCTCTAATGACGGTCTCTACTTCAAGAACGGTGGCAACGAATCTGGTTCTC 543  
QY 122 TACGTTTCTTGAACCTGACAGGCTTTCTGTTATCTTCAATCTCTGCTCTTCTTACAT 181  
Db 542 AAGCTTCTCCAAACCTCAGGTTGCCCATCATCTCTCCCTCTCTCTCTCTCTCTCTCT 483  
QY 182 TACCCGGCGCAGAACCAATGTGGTGAT---AGTACAGTCTTCTTCTTATCAAAACC 238  
Db 482 CACCCGTCACCGTAAACAGAACAAAGAGATGAGAGACTACACTCTTCTCTCATCAACC 423  
QY 239 GCGTCTTCTTATCGCGGCTGTTATTTGTAGGCAATCTCTCAGGGTTTGATAACTTGTGA 298  
Db 422 TCCTCTGTTCTCGGCTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 363  
QY 299 TGCATATGATATAGCTTATCTTCCAGTTTCTACAGCTGCTCTTATCATTTGCTTCAAGTT 358  
Db 362 CGCTTACGGCTTAGCTTACCTCCCGCTCCACCTCTCTCTCTCATCATCTCTCTCTCAGCT 303  
QY 359 AGCTTTTATAGCTATCTTCTCATTTCTCATGTTTAACTAAAGTTTCACTCTCTTTTACCAT 418  
Db 302 AGCCTTACCGGCTCTCTTCGCCCTTCTTATGGTGAAGCAAGAGTTTACGCCCTTCTACTAT 243  
QY 419 CAATGCTGTTGTTGTTGACTGTTGGTGTCTCGGTTTGGGAATGTCATACCGAAACTGA 478  
Db 242 AAACGCGCTGTTTGTCTCACTCTCGGCGCGGAGCGCTTGTCTTTCACGCGGACGGTGA 183  
QY 479 TAAGCCAGTTTATGAGACTCAACAGAGTACATAAATCTGGTTTCTTGTATTTAGTACAGC 538  
Db 182 CAAGCTCCCAAGGAGAGCGCAAGAGTATATAGTTGGGTTTGTATGACGGTGGCTGC 123  
QY 539 AGCTGTTATGATGCTTTTCTATCTTCCCATTTAGTGAATTTGCTTACCAGAAAGCTTAAGCA 598  
Db 122 AGCTGTTCTTACGCGCTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 63  
QY 599 AACCATGAGCTATACCTTGTGCTCGAGTTCCAGTTTGTGTTGTTTGTCTCTCTGTTCTT 656  
Db 62 GAGGATCAGTACACGCTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 5

RESULT 7

CK272858  
 LOCUS EST718936 potato abiotic stress mRNA linear EST 03-AUG-2004  
 DEFINITION clone POAD778 5' end, mRNA sequence.  
 ACCESSION CK272858  
 VERSION CK272858.1 GI:39829836  
 KEYWORDS EST.  
 SOURCE Solanum tuberosum (potato)  
 ORGANISM Solanum tuberosum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 asterids; lamids; Solanales; Solanaceae; Solanum.  
 REFERENCE 1 (bases 1 to 971)  
 AUTHORS Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A. and Baker, B.  
 TITLE Generation of ESTs from abiotic stressed potato tissue  
 JOURNAL Unpublished (2003)  
 COMMENT Other\_ESTs: EST718937  
 Contact: Robin Buell  
 The Institute for Genomic Research  
 9712 Medical Center Dr, Rockville, MD 20850, USA  
 Email: potato-array@tigr.org  
 Clones can be requested from the University of Arizona Genomics  
 Institute via http://genome.arizona.edu/orders/.  
 Seq primer: ATT TAG GTG ACA CTA TAG.

# FEATURES source

1..971  
 /organism="Solanum tuberosum"  
 /mol\_type="mRNA"  
 /cultivar="Kennebec"  
 /db\_xref="taxon:4113"  
 /clone="POAD778"  
 /tissue\_type="abiotic stress treated leaf and root tissue"  
 /lab\_host="DH10B-Tona"  
 /libs="potato abiotic stress cdna library"  
 /note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;  
 supplier: Solanum tuberosum var. Kennebec plants were  
 grown from cuttings on a 16hr light/8 hr dark cycle at 25  
 C for 3-4 weeks. Abiotic stress conditions were applied to  
 four separate sets of plants. Set 1 involved saturation of  
 the soil with 150 mM NaCl and tissues were harvested at  
 following application of the salt stress (leaves: 2hr,  
 6hr, 12hr, 1d, 2d, and 4d; roots: 2hr, 6hr, 12hr, and 2d).  
 Set 2 were grown under the standard conditions and then  
 were water stressed by withdrawal of further watering  
 applications. Drought stressed plants were harvested after  
 cessation of watering (leaves: 3d, 5d, and 7d; roots: 3d  
 and 5d). Set 3 were grown under the standard conditions  
 and then were cold stressed by placement at 4 C. Cold  
 stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,  
 and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d,  
 2d. Set 4 were grown under the standard conditions and  
 then were heat stressed by placement at 35 C. Heat  
 stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,  
 2d and 4d and heat-stressed roots were harvested at 6 hr,  
 12 hr, 1 d, and 4d. RNA was isolated from all tissues and  
 equal RNA from each tissue and stress was pooled to  
 construct the cdna library. RNA sample."

## ORIGIN

Query Match 26.7%; Score 284.4; DB 7; Length 971;  
 Best Local Similarity 58.2%; Pred. No. 1.2e-68;  
 Matches 559; Conservative 0; Mismatches 391; Indels 10; Gaps 3;  
 QY 9 ATGTGGAAGGCTTGTGATCATAACTGCATATTTCTAGCCATAGGAACCTGTGGAGGT 68  
 Db |||||  
 20 ATGAGGAGATCCCTTCGATTAATAATGTTTATTACTATCCGTGGTGGTGGTGGC 79  
 QY 69 CTTTGTATTATCGCTCTACTTCAACAATCGCGTAAAGGATTGTGTTCTCTACCGTTT 128  
 Db |||||  
 80 CTTTAAATGATGCGGTTTATATTTTGTGAAGGTGGTTTAAAGATTATGTTAAATGTTGG 139  
 QY 129 CTTGAACTGCAGGCTTTCCTGTTATCTTCATTCCTCTGCTCTCTTACTATACCCGG 188  
 |||||

Db 140 TTACAAACCGTGGTGGCCACTCACAAATTATACCAATTAGTCATCTTATATTTCTATCGA 199  
 QY 189 CGCAGAGACCAATATGGGGTGATAGTACAAGTTTCTTTCTTATCAAAACCGCTCTCTTT 248  
 Db |||||  
 200 CGAAAAACCAA-----GGGCTCTGATACCAAGTTTATTTTATACACCTCGATTTTC 253  
 QY 249 ATCCCGCTGTTATTATTAGGCATTTCTCTCAGGGTTCATTAACCTTTGTATGTCATATGGT 308  
 Db |||||  
 254 ATTGCATCGTTCAITATTGGGTTTTCACGGGTCTGATTCTTACCTTTATTCAATGGGC 313  
 QY 309 ATAGCTTATCTTCAGTTTCTACAGCTGCTCTTATCAATTTGCTTCACAGTAGTGTATATA 368  
 Db |||||  
 314 GGGTCGAAACTCCCGGTGTCAAACGCTCTCCCTCTCATCGCAGCTCAACTTGCCTTCACG 373  
 QY 369 GCTATCTTCTCATCTTTCATGGTTTAAATAAGTTTCACTCCTTTTACCATCAATGCTGTT 428  
 Db |||||  
 374 CGGATAGGGTCTTACTTTCATAGCAAGATAAATTTTACATCATATTGATTAACCGCGTG 433  
 QY 429 GTGTTGTTGACTGTTGGTGTCTGGGTTTGGGAATGCATACCCGAAATCTGATAAGCCAGTT 488  
 Db |||||  
 434 GTTTTATTGACACTTGGCGGTTTATTGGGTATGCGATCGACGGTGTACGCCGAG 493  
 QY 489 CATGAGACTCACAGCAGTACATACTGGTTCTTGATTTACTGTAGCAGCAGCTGTTATG 548  
 Db |||||  
 494 GGTGTTACAAAGTAAAGAGTATGTTATTGGTTTATTATGACACTTTTGGCTGCACTTTG 553  
 QY 549 TATGCTTTTCATCTTGCCTATTAGTGAACCTTGTCTTACCAGAAAGCTTAAGCAAAACCATGAGC 608  
 Db |||||  
 554 TATGGACTTATTTTACCTGTATTGAGTTGATATATTGGAAGCAAAACAGCTATTACT 613  
 QY 609 TATACCTTGTGCTCGAGTTCCTCAGTTGATTTGTTGCTCTCTCTTCTTATTTGTCAGCGTC 668  
 Db |||||  
 614 GCAACATTGGTGTGGAGATTCAAATGGTCATGGCTTTTGTCTGTCTACTGCTTTTGCACA 673  
 QY 669 ATCGTATGTTTCATCGCTGGTGATTTCAAGCAGCGCTTACCAAAAGAGCAAGAGAGTTC 728  
 Db |||||  
 674 ATTGGAATG---ATTGCTAACAAAGATTTTTCAGGCAATATCAAGGGAAGCAAAACATTTT 730  
 QY 729 AAGCTTCGAGAGGACATGTTCTTATGTGGTGTGTTTTCAGGCATCATATGCAAGGC 788  
 Db |||||  
 731 AACCTTGGAGAGGTAGATATTACATAGTTGTAATATGGAATGCCATTAATTTGCAATGT 790  
 QY 789 TTCTTCTTGGAGCCATTGGATTAAATCTTCCACATCGTCTCTCGTCTCGGSPATTATG 848  
 Db |||||  
 791 TTCTTTGTGCTGTGGAGTTATTTACTCTCTCTTCTTTTAAATGTCGTGAGTTATG 850  
 QY 849 ATATCAGTGTCTTTGCCAATTCAGAGGTTTATAGCTGTTATATCTACCATGAAGATTT 908  
 Db |||||  
 851 ATTGCAGTTTCTTACGCCAATTTACTGAAAGTATTAGGAGTAATTTCTTTGATGAAAAATCT 910  
 QY 909 CAAGCTCAGAAAGGACATTTCTCTTGTCTCTCCCTTTGGGGCTTTGTCTCTTACTTTTAT 968  
 Db |||||  
 911 C-AGCTGAAAAGGACATTTTCACTTTTCTTCTTGTGGGTTTGTCTTTCTTATTTT 969

RESULT 8  
 CK243751  
 LOCUS CK243751  
 DEFINITION EST727388 potato callus cdna library, normalized and full-length  
 Solanum tuberosum cdna clone POCAS40 5' end, mRNA sequence.  
 ACCESSION CK243751  
 VERSION CK243751.1 GI:39788661  
 KEYWORDS EST.  
 SOURCE Solanum tuberosum (potato)  
 ORGANISM Solanum tuberosum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 asterids; lamids; Solanales; Solanaceae; Solanum.  
 REFERENCE 1 (bases 1 to 986)  
 AUTHORS Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A. and Baker, B.  
 TITLE Generation of ESTs from potato callus tissue  
 JOURNAL Unpublished (2003)  
 COMMENT Contact: Robin Buell



```
QY 145 TTCCTGTTATCTTCATTCCTCTGCTCTCTCTTACATTACCGCGCGAGAACACAAATG 204
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 176 GTCCAAATTAATCTCAATCCCTAGCGATAGCCTATTTCAACGCCGGAATTTCAAGAC 235
QY 205 TGGGTGATAGTACAAGTTTCTTTTCTTATCAACCGCGTCTTCTTATCGCCGCTGTATTG 264
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 236 CCGAGGCTAAAA-----TCGCTTTATCACAGCGCAAGAGTTTCATCGCATCTGCCGCG 289
QY 265 TAGGCATTTCTCAGGGTTTGATACTACTTGTATGATGATATGATAGCTATCTTCCAG 324
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 290 TCGGAATCATCGTCGGTCTCGATGTTTACTTAAACTCATGGGAGCCCGGAAATTAACCG 349
QY 325 TTTCACAGCTGCTCTTATCATTTGCTTCACAGTTAGCTTTTATAGCTATCTTCATCT 384
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 350 TTTCACACTTCGACATAATCAACGGCACTCAACTTGGCTTCACTCGGCTTTTGTGTGC 409
QY 385 TCATGGTTAAACATAAAGTTCACTCTCTTTTACCATCAATGCTGTTGTGTTGACTGTG 444
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 410 TTATAGTTAAACAGAAATTTGACGGTATTCGACGAATTCGGTCGTTTGTCTTATCGCG 469
QY 445 GTGCTGCCGTTTGGGAATGCATACCGAACTGATAGCCAGTTTCATGAGACTCACAAAG 504
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 470 GAGCTGCCACTTTAGCTCTCCGGCGGAACGGTGACCGCGCGCGCGAGTCCAGGAG 529
QY 505 AGTACATAACTGGTTTCTTGATTACTGTAGCAGCAGCTGTTATGATGCTTTTCATCTTGC 564
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 530 ATTATATGTTGGGTTTGTGATGAGCTTTATCGGTGCGGTGTTATATGATGATGTTGC 589
QY 565 CATTAGTGGAACTTGTCTTACCAGAAAGTCAAGCAACCATGAGCTATACCCCTTGTGCTCG 624
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 590 CGTTAAATGAGTTGATTTATATGAGGCGGAAGAGCTGTTACTTTACACTACAGTATTGG 649
QY 625 AGTTCAGTTGATTTGTGTCCTTGTCTTATGTCAGCGTCAATCGGTATGTTCAFCG 684
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 650 AGATTACAGATGGTTTGGGCAATTTCTGCTACTGTTTGTGCACTATTGGAAATGGTTATCA 709
QY 685 CTGTGATTTTCAAGCAGCGCTTACCAGAAAGCAAGAGAGTTCAAGCTTCGAGAGGCAT 744
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 710 ACAGGATTTCT---CAGCGGATTCAGAGGAGCAAGTCAATATGCAATGGAGAACTA 766
QY 745 TGTCTATGTGGTGCTGTGTTTACGCCATCATATGGCAAGGCTTCTTCTTGGGAGCCA 804
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 767 AGTATTATATGATAGTATGTTGGTGTGCAATATTTGGCAATTCGCGCTCCTGGGTTAG 826
QY 805 TTGGATTAATCTTCCACATCGTCTCGCTCTCGGTATTTATGATATCATGTCGTTTTCG 864
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 827 TTGGAGTTAATTTTATTTCTTCTCTTACTCTCTGGAATTTATAGGCGCTTTTACTTTC 886
QY 865 CAATTACAGAGTTTATAGCTGTTATATTTCTACCATGAAAGTTTCAAGCTGAGAGGAC 924
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 887 CTGTTACCGAGTTTGGCTGTAATTTGTTCCAGGAAATTTCAAGCCGAAAAAGAG 946
QY 925 TTTCCTGTGCTCTCCCTTTGGGGCTTTGTCCTTACT 963
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 947 TAGCTAATTTTCTAGCTCTATGGGATTTGTTTCATATT 985
```

```
RESULT 10
AJ807511 LOCUS
DEFINITION AJ807511 Antirrhinum majus whole plant Antirrhinum majus cDNA clone
            018.6.06.p14, mRNA sequence.
ACCESSION AJ807511
VERSION AJ807511.1 GI:51122839
SOURCE EST.
ORGANISM Antirrhinum majus (snapdragon)
           Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
           Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
           asterids; lamiales; Lamiales; Plantaginaceae; Antirrhineae;
           Antirrhinum.
REFERENCE 1 (bases 1 to 688)
```

AUTHORS  
TITLE  
JOURNAL  
COMMENT

Zachgo, S., Stueber, K., Saedler, H., Sommer, H. and Schwarz-Sommer, Z.  
Antirrhinum EST collection  
Unpublished (2003)  
Contact: Schwarz-Sommer Z  
Molekulare Pflanzengenetik  
MPI fuer Zuechtungsforschung  
Carl-von-Linne Weg 10, D-50829, Germany.  
Location/Qualifiers

FEATURES  
source

1..688  
/organism="Antirrhinum majus"  
/mol\_type="mRNA"  
/db\_xref="taxon:4151"  
/clone="018.6.06.p14"  
/tissue\_type="whole plant"  
/clone\_lib="Antirrhinum majus whole plant"

## ORIGIN

Query Match 24.9%; Score 265.2; DB 1; Length 688;  
Best Local Similarity 64.0%; Pred. No. 2.7e-63;  
Matches 432; Conservative 0; Mismatches 239; Indels 4; Gaps 2;  
QY 330 ACAGCTGCTCTTATCATGTGTTACAGTTAGCTTTTATAGCTATCTTCTCATTTCTCATG 389  
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 11 ACCTCGTCACTGATTATTGCGACGCTGTCGCTTACTGCGGGTTTGCCTTATTTCTT 70  
QY 390 GTTAACATAAAGTTCACCTCTTTTACCATCAATGCTGTTGTGTTGACTGTTGGTGCT 449  
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 71 GTGAAGCAGAGATTTCACGGCGTTT-TNACAAACGATCGTTTTGTGCTGACGTTGGGAGCG 129  
QY 450 GCGGTTTTGGGAATGCATACCGAAACTGATAAGCCAGTTTCATGAGACTCACAAAGCAGTAC 509  
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 130 GTGGTTCTAGGGCTGACACACAGCGCGATAGGCCGAGGAGAGTCAACAAAGAGTAT 189  
QY 510 ATAACTGGTTTCTGATTACTGTAGCAGCAGCTGTTATGATGCTTTTCACTTGGCCATTA 569  
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 190 TTAGCTGGGTTTTTATGACGCTTGGGCGCGCGCGCTTATGCGTTTATTTTGGCATTTG 249  
QY 570 GTGGAACCTCTCTTACCAGAAAGCTAAGCAAAACCATGAGCTATACCCCTGTCGAGTTC 629  
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 250 GTGGAGTTGACGTATATGAAGCGGAGGAGCTTGGATTACACACTCGTATTGGAAAT 309  
QY 630 CAGTTGATTTTGTCTCTCTCTGCTTCTTATTTGTGACGCTCATCGGTATGTTTCATCGCTGGT 689  
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 310 CAGTTGCTCATGTGTTCTTCGCACTGCTTTTGTACTGTTGGAATGCTCGTCAACAA 369  
QY 690 GATTTCAAGCAGGCTTACCAAAGAAAGCAAGAGAGTTCAAGCTTGGAGAGGCAATGTTTC 749  
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 370 GATTTCT---CAGGCAATTCCAAGAGCAAGAGCGTACGAGCTAGGGGAGACAAAATAC 426  
QY 750 TATGCTGGGCTGTGTTTTTCAGCCATCATATGGCAAGGCTTCTCTTGGGAGCCATTGGA 809  
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 427 TATTAGTGTAGTGTGTTTAGTGCCATAAATTTGGCAGTGTCTTTCTTGGGAGCAATTTGGA 486  
QY 810 TTAATCTTTCCACATCGTCTCTCGTCTCGGGTATTATGATATCAGTGTCTTTTGGCCAAAT 869  
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 487 ATCATATTTTACTCTTCGCTCTCTTTTATCGGCAATCGTAATTAATTAATTAATTAATTA 546  
QY 870 ACAGAGGTTTTAGCTGTATTAATTTACCATGAAAGTTTCAAGCTGAGAGGGAATTTCT 929  
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 547 ACAGAAATTTGTTGGCTGTATTTTATTTTACCATGAAATAATTTCAAGCAGAAAAAGGTTATTTCT 606  
QY 930 CTTGCTCTCTCCCTTTGGGCGCTTGTCTCTTACTTTTATGCTGAGATAAAGTCTGGCGAG 989  
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 607 CTTTCTCTCTCGGTGTGGGATTTTGTCCGTACTTTTATGTTGAGATCAAAACACTTGAAG 666  
QY 990 GATAAAAGGAGAAAT 1004  
Db 667 AAGATGAAGGGAAGT 681

RESULT 11  
CK251896  
LOCUS

CK251896 1013 bp mRNA linear EST 30-JUL-2004



```

DEFINITION EST735533 potato callus cDNA library, normalized and full-length
ACCESSION CK251896
VERSION CK251896.1 GI:39805365
KEYWORDS EST.
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum.
1 (bases 1 to 1013)
Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A. and Baker, B.
Generation of ESTs from potato callus tissue
Unpublished (2003)
Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@igr.org
Clones can be requested from the University of Arizona Genomics
Institute via http://genome.arizona.edu/orders/
Seq primer: ATT TAG GTG ACA CTA TAG.
Location/Qualifiers
1..1013
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="POC8R03"
/tissue_type="callus"
/lab_host="DH10B-Tona"
/clone_lib="potato callus cDNA library, normalized and
full-length"
/note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;
supplier: RNA was isolated from Solanum tuberosum var.
Kennebec callus tissue grown on solid media."

FEATURES
source
1..1013
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="POC8R03"
/tissue_type="callus"
/lab_host="DH10B-Tona"
/clone_lib="potato callus cDNA library, normalized and
full-length"
/note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;
supplier: RNA was isolated from Solanum tuberosum var.
Kennebec callus tissue grown on solid media."

ORIGIN
Query Match 24.8%; Score 264.8; DB 7; Length 1013;
Best Local Similarity 57.2%; Pred. No. 4e-63;
Matches 543; Conservative 0; Mismatches 397; Indels 10; Gaps 3;

QY 25 TCATCAATCACTGCTATCTAGCATAGGAACCTGGAGGTCCTTTGATATGCGTC 84
Db 63 TAAATTTCAACATAATATGCTTTTCAATCGGAATTCGGTGTCTCTTAATCTCTCGAC 122
QY 85 TCTACTTCAACATGCGGTGAAAGATTGGTTCTCTACGTTTCTTGAAGCTGCGAGCT 144
Db 123 TCTATTTCAATTCATGGCGCCACGAAATTTGATACCCAGCGATTAACAACCTGTGGTT 182
QY 145 TTCTCTGTTATCTTCAATCTCTGCTCTCTTCTTACATTTACCGCGGCGAAGCAACAATG 204
Db 183 GTCCAAATATTCTCATCCCTCTAGCATAGCCTATTTCACGCGGAAATTCAGGAC 242
QY 205 TGGGTGATAGTACAAGTTCTTTCTTATCAACCGGCTCTTATTCGCGCTGTATTG 264
Db 243 CCGAGGCTAAAA-----TCGCTTTATCACGCGCAAGATTCATCGCATCTGCGGCG 296
QY 265 TAGGCATCTCTCAGGGTTTGATACTACTTGTATGCATATGATAGCTTATCTTCCAG 324
Db 297 TCGGAATCATCTGCTCTCGAATGTTTAACTATGCGGACCCGCGAAATACCCG 356
QY 325 TTTCTACAGCTGCTCTTATCATTTGCTTTCACAGTTAGCTTTTATAGCTATCTTCTCATTTCT 384
Db 357 TTTCAACTTCGACATTAATCAACGCGACTCACTTGGTTCCTACTGGCTTTTGTGTGTC 416
QY 385 TCATGGTTAAACATAAAGTTCACTCTCTTTTACATCAATGCTGTTGTGTGTTGACTGTG 444
Db 417 TTATAGTTAAACAGAAATGACAGGCTATTTCGACGAATTCCTGCTGTTTGTCTTATCGCG 476
QY 445 GTGCTGCGGTTTTGGGAATGCATACCGAACTGATAAGCCAGTTTATGAGACTCAAGC 504
Db 477 GAGCTGCGCACTTAGCTCTCCGGGCGAAGCGGTGACCGGCGCCGCGGAGTCCAGGAG 536

QY 505 AGTACATAAATCGGTTTCTTGATTTACTGTAGCAGCAGCTGTTATGATGCTTTTCATCTTGC 564
Db 537 ATTATATGTTGGGTTTGTGATGACGTTTATCGGTGCGGTGTTATATGATGATGCTTGC 596
QY 565 CATTAAGTGAACCTTGCCTTACCAAGAAAGCTTAAGCAACCAATGAGCTATACCTTTGCTCG 624
Db 597 COTTAATTTGAGTTGATTTATATGAAGGCGAAGCAAGCTGTTACTTACACTACAGTATTGC 656
QY 625 AGTTCCAGTTGATTTTGTGTCCTTCTTCTTATTTGTGAGCTCATCGGTATGTTTCATCG 684
Db 657 AGATTCAGATGTTTGGGCAATTTCTGCTACTGTTTGTGCACTATTTGGAATGTTATCA 716
QY 685 CTGTGTGATTTCAAGCAGCGCTTACCAAGAAAGCAAGAGTTCAGCTTGGAGAGGCAT 744
Db 717 ACAAGGATTTCC---CAGGCGATTTCAAGGGAAGCAAGTCAATATGCAATTTGGAGAATCTA 773
QY 745 TGTTCTATGTTGGTGTGCTGTTTTCAGGCATCATATGCAAGGCTTCTTCTTGGAGCCA 804
Db 774 AGTATATATGTTGCTACTAGTATGTTGCAATTTATTTGGCAATTCGCGCTCTCTGGGTTAG 833
QY 805 TTGGATTAATCTTCTCCACATCGTCTCTGCTCGGTGTTATGATATATCAGTCTTTTTC 864
Db 834 TTGGAGTTATTTTATTTCTTCTTCTTCTCTGGAATTTATAGGCGCTTTTACTTCTC 893
QY 865 CAATTCAGAGGTTTGTAGCTGTTATTTCTACCATG-AAAAAGTTTCAAGCTGAGAGGGA 923
Db 894 CTGTTTACGAAGTTTGGCTGTAAATTTTGTTCACGAAAAAATTTCAAGCCGAAAAAGGA 953
QY 924 CTTTCTCTGCTCTCTCCCTTGGGCTTGTGCTCTTACTTCTTATGTTGTTGA 973
Db 954 GTAGCTATTTTCTAGCTCTATGGGATTTGTTTCATATTTTATGTTGA 1003

RESULT 12
CK247918
LOCUS CK247918
DEFINITION Solanum tuberosum cDNA clone POCAX15 5' end, mRNA sequence.
ACCESSION CK247918
VERSION CK247918.1 GI:39797543
KEYWORDS EST.
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum.
1 (bases 1 to 998)
Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A. and Baker, B.
Generation of ESTs from potato callus tissue
Unpublished (2003)
Other ESTs: EST731556
Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@igr.org
Clones can be requested from the University of Arizona Genomics
Institute via http://genome.arizona.edu/orders/
Seq primer: ATT TAG GTG ACA CTA TAG.
Location/Qualifiers
1..998
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="POCAX15"
/tissue_type="callus"
/lab_host="DH10B-Tona"
/clone_lib="potato callus cDNA library, normalized and
full-length"
/note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;
supplier: RNA was isolated from Solanum tuberosum var.
Kennebec callus tissue grown on solid media."

FEATURES
source
1..998
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="POCAX15"
/tissue_type="callus"
/lab_host="DH10B-Tona"
/clone_lib="potato callus cDNA library, normalized and
full-length"
/note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;
supplier: RNA was isolated from Solanum tuberosum var.
Kennebec callus tissue grown on solid media."

ORIGIN

```

Query Match	24.6%	Score 262.8	DB 7	Length 998
Best Local Similarity	56.5%	Pred. No. 1.5e-62		
Matches 533	Conservative 0	Mismatches 402	Indels 9	Gaps 2
Qy	25	TGATCATAAATCGCATAACTTAGCCATAGGAARACTGGAGTCCCTTGGATTATGGTC	84	
Db	48	TAAATTTTCAACATAAATATGCTTTCAATCGGCAATGGGTGGTCCCTTAACTCTCGAC	107	
Qy	85	TCTACTTCAACAATGGGGGAAAAAGGATTTGGTTCTCTACGTTTCTTGAACCTCAGGCT	144	
Db	108	TCATTTTCAATTCATGGCGGCAACGAATTTGGATACCCAGCGTATTACAACTCTGGTT	167	
Qy	145	TTCTGTGTAATCTCATCTCTCTGCTCTTCTCTTACATTAACCGGGGAGAAACAATG	204	
Db	168	GTCCAAATTAATCTCATCCCTCTAGCCATAGGCTATTTTCCAACGCGGAAATTTCAAGGAC	227	
Qy	205	TGGGTGATAGTACAAGTTCTTCTTATCAAAACCGGCTCTCTTATCGCGCTGTATTG	264	
Db	228	CCGAGGCTAAAA-----TGGTCTTTTATCACACGCAAGTTTCATCGATCTCCGGCG	281	
Qy	265	TAGGCATTTCTCAGGGTTTGATAACTACTTGTATGCATATGGTATAGCTTTATCTTCAG	324	
Db	282	TCGGAATCATCGTCGGTCTCGATGGTTACTTTAAACTCATGGGACCCGCGAAATTTACCG	341	
Qy	325	TTTCTACAGCTGCTTATCATTCGTTTCAAGTTAGCTTTTATAGCTATCTTCTCATCT	384	
Db	342	TTTCAACTTTCGACACTAATCAACCGCACTCAACTTGGGTTCACTGCGCTTTTTCGTGTC	401	
Qy	385	TCATGGTTTAACATAAAGTTCACTCCTTTTACCATAATGCTGTGTGTGTGTGACTGTG	444	
Db	402	TTATAGTTTAAACAGAAATTTGACAGGTTATTCGACGAATTCGTCGTTTTCGTTATCGCG	461	
Qy	445	GTGCTCGGTTTTGGGAATGCATACCGAAACTGATAAGCCAGTTTCATGAGACTCACAAGC	504	
Db	462	GAGCTCGCACTTTAGCTCTCGGCGCAACGGTGAACGCGCGCGCGGAGTCCACGAAGG	521	
Qy	505	AGTACATAACTGGTTCTTGAATTACTGTAGCAGCAGCTGTATGTAATGCTTTTCACTTGC	564	
Db	522	ATTATATGTTGGGGTTTGTGAATGAGCTTTATCGGTGCGGTGTTATATGGATTTGATGTTG	581	
Qy	565	CATTAGTGGAACTTGTCTTACAGAAAGCTTAAGCAAAACCATGAGCTATACCCCTTGTCGTC	624	
Db	582	CGTTAATTGAGTTGATTTATATGAAGGGAAGCAAGCTGTACTTACACTACAGTATTGG	641	
Qy	625	AGTTCCAGTTGATTTTGTGTCCTCTGTTCTATGTTCAGGCTCATCGGTATGTTCAATCG	684	
Db	642	AGATTCAAGTGGTTTTGGGCATTTCTGCTACTGTTTTTTTCGACTATTTGGAAATGGTTATCA	701	
Qy	685	CTGGTGATTTCAACAGCGCCTTACCAAAAGAAGCAAGAGAGTTTCAAGCTTGGAGAGGCAT	744	
Db	702	ACAAGGATTTCT---CAGCGAATCCAGGGAAGCAAGTCAATATGCAATGGAGAACTTA	758	
Qy	745	TGTTCTATGTGGTGGCTGTGTTTTTTCAGCCATCATATGGCAAGGCTTCTTCTTGGGAGCCA	804	
Db	759	AGTATTATATGTTACTAGTATGGTGTGCCATTTATTTGGCAATTCGCGCTCTCGGGTTAG	818	
Qy	805	TTGGAATTAATCTTCCACATCGTCTCTGTCCTGGGTATTTATGATATCAGTGTCTTTTGC	864	
Db	819	TTGAGGTATTTTTTATTTCTCTCTTACTCTCTCGAATTTATAGGCGCTTTTTTACTTTC	878	
Qy	865	CAATTCACAGAGTTTTAGCTGTATATTTCTACCAATGAAAAGTTTCAAGCTGAGAGGAC	924	
Db	879	CTGTTACCGAAGTTTTGGCTGTAAATTTGTTCCAGCAAAATTTCAAGCGCAAAAAGAG	938	
Qy	925	TTTCTCTTGTCTCTCCCTTTGGGCTTTGTCTCTTACTTTTAT	968	
Db	939	TAGTATTTTTCTAGCTCTATGGGGATTTGTTTCATATTTTTT	982	

RESULT 13  
CK257887  
LOCUS

```

DEFINITION EST741524 potato callus cDNA library, normalized and full-length
Solanium tuberosum cDNA clone POCDB85 5' end, mRNA sequence.

ACCESSION CK257887
VERSION CK257887.1 GI:39814867
KEYWORDS EST.
SOURCE Solanium tuberosum (potato)
ORGANISM Solanium tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum.
REFERENCE 1 (bases 1 to 907)
AUTHORS Buell,C.R., Hart,A., Zismaun,V., Karamycheva,S.A. and Baker,B.
TITLE Generation of ESTs from potato callus tissue
JOURNAL Unpublished (2003)
COMMENT Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@tigr.org
Clones can be requested from the University of Arizona Genomics
Institute via http://genome.arizona.edu/orders/
Seq primer: ATT TAG GTG ACA CTA TAG.

FEATURES             Location/Qualifiers
     source            1..907
                        /organism="Solanum tuberosum"
                        /mol_type="mRNA"
                        /cultivar="Kennebec"
                        /db_xref="taxon:4113"
                        /clone="POCDB85"
                        /tissue_type="callus"
                        /lab_host="DHL0B-Roma"
                        /clone_lib="potato callus cDNA library, normalized and
                        full-length"
                        /note="vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;
                        supplier: RNA was isolated from Solanium tuberosum var.
                        Kennebec callus tissue grown on solid media."

```

[illegible]

```
QY 550 ATGCTTTTCATCTTCCCATTAAGTGAACCTGCTTACACGAAAGCTAAGCAAAACCATGAGCT 609
Db 478 ATGGATTGATGTTGCCGTTAAATGAGTTGATTTATATGAAGCGCAAGCTGTTACTT 537
QY 610 ATACCTTGTGCTCGAGTTCAGTTGATTTTGTGCTCCTTGTCTTATTTGTCAGGCTCA 669
Db 538 ACACTACAGATTGGAGATTCAATGGTTTGGGCAATTTCTGCTACTGTTTTTGGCACTA 597
QY 670 TCGGTATGTTTATCGCTGTGATTTCAAGCAGGCTTTACCAAGAGCAAGCAAGAGATTCA 729
Db 598 TTGGATGATTATCAACAAGATTTC--CAGCGATTCCAAAGGAAGCAAGTCAATATG 654
QY 730 AGCTTGAGAGGCAATGTTCTATGTGCTGTGCTGTTTTTTCAGCCATCATATGCAAGGCT 789
Db 655 CAATTGGAGAATCTAAGTATTATATGTAAGTATGTTGCTGCTGCTGCTGCTGCTGCTG 714
QY 790 TCTTTCTGGGAGCAATGGAATTAATCTTCCACATGCTCTCGTCTGGGTATTATGA 849
Db 715 CGCTCTGGGTTAGTTGGAGTTATTTTATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 774
QY 850 TATCAGTGTCTTTTGGCAATTACAGAGTTTGTAGCTGTATATATTTCTACCATGAAAGTTTC 909
Db 775 GCGCTTTTATCTTCTGTTACCGAAGTTTGGCTGTATTTGTTCCAGCAAAATTTTC 834
QY 910 AAGCTGAGAGGACTTTCTTGTGCTCTCTCCCTTTGGGGCTTTGTCTCTTACTTTATG 969
Db 835 AAGCCGAAAGAGTAGCTATTTTCTAGCTCTATGGGGAATTTGTTTCATATTTTATG 894
QY 970 GTGA 973
Db 895 GTGA 998
```

```
RESULT 14
CK248243
LOCUS
DEFINITION
EST731880 potato callus cDNA library, normalized and full-length
Solanum tuberosum cDNA clone PCAZ44 5' end, mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Solanum tuberosum (potato)
Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum.
REFERENCE
1 (bases 1 to 988)
Buell,C.R., Hart,A., Zismann,V., Karamycheva,S.A. and Baker,B.
Generation of ESTs from potato callus tissue
Unpublished (2003)
Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@igr.org
Clones can be requested from the University of Arizona Genomics
Institute via http://genome.arizona.edu/orders/
Seq primer: ATT TAG GTG ACA CTA TAG.
FEATURES
Location/Qualifiers
1..998
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="PCAZ44"
/tissue_type="callus"
/lab_host="DH10B-Tona"
/clone_lib="potato callus cDNA library, normalized and
full-length"
/notes="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;
supplier: RNA was isolated from Solanum tuberosum var.
Kennebec callus tissue grown on solid media."
```

ORIGIN

```
Query Match 23.1%; Score 246.6; DB 7; Length 988;
Best Local Similarity 56.2%; Pred. No. 5.4e-58;
Matches 527; Conservative 0; Mismatches 399; Indels 11; Gaps 3;
QY 25 TCATCATAACTGCATAATTTCTAGCCATAGGAACCTGTGGAGGTCTTTGATTATCGGTC 84
Db 62 TAAATTTTCAACATAATATATGCTTTCAATCGGCAATTTGCGGTGGTCTCTTAACTCTCTCGAC 121
QY 85 TCTACTTCAACAATGCGGTAAGGATTTGGTTCTCTACGTTTCTTGAACATGCAAGGCT 144
Db 122 TCTATTTTCAATGCGGCAAGGATTTGATATACCCAGCGATTTACAACTGTTGGTT 181
QY 145 TTCTCTGTTATCTCAATTCCTCTGCTCTCTTCAATATACCCGGGCGAGAACAAATG 204
Db 182 GTCCAAATTTCTATCTCTCTAGCATAGCCCTATTTCCAAAGCGGAAATTTCAAGGAC 241
QY 205 TGGGTGATAGTACAAGTTTCTTCTTATCAAAACCGGCTCTTCTATCGCGGCTGTTATG 264
Db 242 CCGAGGCTAAAA-----TCGCTTTTATCACAGCGCAAGAGTTTCATCGCATCTGCGGCG 295
QY 265 TAGGCATTTCTCAGGTTTGTATTAATACTTCTGTATGCATATGTTATGCTTATCTTCCAG 324
Db 296 TCGGAATCATCTGCTCGTCTCGATGTTTAACTCATGGGACCCGCGAATTTACCCG 355
QY 325 TTCTTACAGCTGCTCTTATCATTTGCTTCAAGTTAGCTTTTATAGCTATCTTCTCATTTCT 384
Db 356 TTTCAACTTCGACTCTAATCAACGCGACTCAACTTGGCTTCTACTGCGCTTTTGTCTGTGC 415
QY 385 TCATGTTTAAACATAAAGTTTCACTTCTTTTACCATCAATGCTTGTGTTGTTGAGCTGTTG 444
Db 416 TTATAGTTTAAACAGAAATTTGACAGCGTATTCGACGAAATTCGCTGCTTTTCTTATCGCG 475
QY 445 GTGCTGCGGTTTGGGAATGTCATACCGAAACTGATAAGCCAGTTTCATGAGACTCAACAAG 504
Db 476 GAGCTGCGACTTTAGCTCTCCGGGCGAAGCGGTACCGCGCGCGCGAGTTCACGAAAG 535
QY 505 AGTACATAACTGTTTTTCTTGTACTGTAGCAGAGCTGTTATGATGCTTTTCACTCTTCG 564
Db 536 ATTATATGTTGGGTTTGTGATGACGTTTATCGGTGCGGTGTTATATGATTTGATGTTGC 595
QY 565 CATTAGTGGAACTTCTTACCAGAAAGCTTACCAAAACCATGAGCTATACCTTTGCTCTG 624
Db 596 CGTTAATTGAGTTGATTTATATGAAGCGCAAGCAAGCTGTTACTTACACTACAGTATTGG 655
QY 625 AGTTCAGTTGATTTTGTCTCTCTTCTTATGTCAGCTCATCGGTATGTTTCATCG 684
Db 656 AGATTACAGATGTTTTGGGCAATTTCTGCTACTGTTTTTGGCACTATTTGAATGATTATCA 715
QY 685 CTGGTGATTTCAAGCAGGCTTTACCAAGAGCAAGAGAGTTCAAGCTTTGGAGAGGCAT 744
Db 716 ACAAGGATTTTC--CAGGCGATTCCAAAGGAAGCAAGTCAATATGCAATTTGGAGAATCTA 772
QY 745 TGTCTATGTTGGTGGCTGTTTTCAGGCATCATATGGAAGGCTTCTCTTTGGGAGCCA 804
Db 773 AGTATTATATGTTACTAGTATGTTGCTGCTATTTATTTGGCAATTCGCGCTCTCTGGGTTAG 832
QY 805 TTGGATTAATCTTCTCCACATCTCTCTGCTCTCGGTTATTTATATATCATCAGTCTTTTTCG 864
Db 833 TTGGAGTTATTTTTTATTTCTTCTCTTCTTCTCTCTGGAATTTATAGGCGCTTTTACTTCT 892
QY 865 CAATTACAGAGGTTTTTGTGTTATTTCTTACCATGAAAGTTTCAAGCTGAGAGGAGAC 924
Db 893 CTGTTTACGAGCTTTTGGCTGTAAATTTTGTTCACGAAAAA--TTTCAGCGCGAAAAAGGAG 950
QY 925 TTTCTCTTGTCTCTCTCCCTTTGGGCTTTGTCTCTTA 961
Db 951 TAGCTATTTTCTAGCTCTATGGGATTTGTTTTCATA 987
```

RESULT 15  
CK252005  
LOCUS  
DEFINITION  
CK252005  
EST735642 potato callus cDNA library, normalized and full-length

Solanum tuberosum cDNA clone POCBR86 5' end, mRNA sequence.  
CK252005  
VERSION CK252005.1 GI:39805578  
KEYWORDS EST,  
SOURCE Solanum tuberosum (potato)  
ORGANISM Solanum tuberosum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; lamids; Solanales; Solanaceae; Solanum.  
1 (bases 1 to 979)  
REFERENCE Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A. and Baker, B.  
AUTHORS Generation of ESTs from potato callus tissue  
TITLE Unpublished (2003)  
JOURNAL Other ESTs: EST735643  
COMMENT Contact: Robin Buell  
The Institute for Genomic Research  
9712 Medical Center Dr, Rockville, MD 20850, USA  
Email: potato-array@tigr.org  
Clones can be requested from the University of Arizona Genomics  
Institute via <http://genome.arizona.edu/orders/>.  
Seq primer: ATT TAG GTG ACA CTA TAG.

FEATURES  
source  
1..979  
/organism="Solanum tuberosum"  
/mol\_type="mRNA"  
/cultivar="Kennebec"  
/db\_xref="taxon:4113"  
/clone="POCBR86"  
/tissue\_type="callus"  
/lab\_host="DH10B-TonA"  
/clone\_lib="potato callus cDNA library, normalized and  
full-length"  
/note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;  
supplier: RNA was isolated from Solanum tuberosum var.  
Kennebec callus tissue grown on solid media."

## ORIGIN

Query Match 23.1%; Score 246; DB 7; Length 979;  
Best Local Similarity 56.4%; Pred No. 8e-58;  
Matches 523; Conservative 0; Mismatches 395; Indels 10; Gaps 3;  
QY 25 TGATCATAACTGCATAATCTTAGCCATAGGAAACTGTGGAGGTCTTTGATTATGCGTC 84  
Db 62 TAAATTTCAACATAATATATGCTTTCAATCGGCAATTTGCGGTGCTTTAATCTCTCGAC 121  
QY 85 TCTACTTCAACATGCGGGTAAAGGATTTGGTCTCTACGTTCTTGAAACTCGAGGCT 144  
Db 122 TCTATTTCAATGCGGCGCAACGAATTTGGATACCCAGCGTATTAACAACTGTGGTT 181  
QY 145 TTCCCTGTTATCTTCAATCTCTGCTCTCTTACATTTACCGGCGCAGAACCAATG 204  
Db 182 GTCCAAATTAATCTCATCTCTCTAGCCATAGCTATTTCCAGCGCGGAAATTTCAAGGAC 241  
QY 205 TGGGTGATGATGACAAGTTCTTTCTTATCAACCGCGTCTTTTATCGCCGCTGTATTG 264  
Db 242 CCGAGGCTAAAA-----TGGTCTTTATCACAGGCAAGAGTTTCATCGCATCTGCGGCG 295  
QY 265 TAGGCATTTCTCAGGGTTTGATTAATACTATGTTGATGATATGATAGCTTATCTTCAG 324  
Db 296 TCGGAATCATCGTCGGTCTCGATGGTTTACTTAAACTCATGGGAGCCCGCGAAATTTACCCG 355  
QY 325 TTTCCTACAGCTGCTCTTATCATTTGCTTCACAGTTAGCTTTTATAGCTATCTTCTCATTTCT 384  
Db 356 TTTCAACTTCGACTATTAACAGCGACTCAACTTGGCTTCAGTGGCTTTTGTGCTGTC 415  
QY 385 TCATGGTTAACAATAAAGTTCACTCTCTTTTACATCAATGCTGTTGTGTTGATGCTGTG 444  
Db 416 TTATAGTTTAAACAGAAAATTGACAGGCTATTCGACGAATTCGCTCGTTTGTCTATCGCGC 475  
QY 445 GTGCTGCGGTTTTGGGAATGCATACCGAAACTGATAAGCCAGTTTCATGAGACTCACAGC 504  
Db 476 GAGCTGCGACTTTAGCTCTCCGGGCGAAACGGTGACCGCGCGCGCGAGTCCACGAAG 535

QY 505 AGTACATAAATCGGTTTCTTGATTACTGTAGCAGCAGCTGTTATGTATGTATGTCTTCACTTGC 564  
Db 536 ATTATATGTTGGGCTTTGTGATGACGTTTATCGGTGCGGTGTTATATGGATTGATGTTGC 595  
QY 565 CATTAGTGGAACTTGTCTTACAGAAAGCTAAGCAACCATGAGCTATACCCCTTGTGCTCG 624  
Db 596 CGTTAAATTGAGTTGATTATATGAAGCGGAAGCAAGCTGTTACTTACACTACAGTATTGG 655  
QY 625 AGTTCCAGTTGATTTTCTGCTCTCCTTGTCTTATTTGTACGCTCATCGGTATGTTTCATCG 684  
Db 656 AGATTCAGATGGTTTTGGGCATTTCTGCTACTGTTTTTTCGACTATTGGAAATGATTATCA 715  
QY 685 CTGCTGATTTCAAGCAGGCTTACCAAAGAAAGCAAGAGAGTTTCAAGCTTTGGAGAGCAT 744  
Db 716 ACAAGGATTTTC---CAGGCGATTTCCAAGGGAAGCAAGTCAATATGCAATTTGGAAATCTA 772  
QY 745 TGTTCTATGTGGTGGCTGTGTTTTCAGCCATCATATGGCAAGGCTTCTTCTTGGAGGCCA 804  
Db 773 AGTATTATATGGTACTAGTATGTTGTCCTATTTTGGCAATTTGCGCTCTCTGGGTTAG 832  
QY 805 TTGGATTAATCTTCTCCACATCGTCTCTCGTCTCGGCTATTATGATATCATGCTGCTTTTC 864  
Db 833 TTGGAGTTATTTTATTTCTTCTCTTCTCTGGAATTTAGGCGCTTTTTCATCTC 892  
QY 865 CAATTACAGAGGTTTTCAGCTGTTTATATTTTACCATGAAAAGTTTTCAGCTGAGAGGAC 924  
Db 893 CTGTTTACGAAAGTTTGGCTGTAAT-TTTTTTCAAGCAAAATTTCAAGCCGAAAAAGGAG 951  
QY 925 TTTCTCTTGTCTCTCCCTTTGGGCTT 952  
Db 952 TAGCTATTTTCTAGCTCTATGGGATT 979

Search completed: November 1, 2004, 22:30:29  
Job time : 3805.29 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 1, 2004, 06:20:04 ; Search time 552.64 Seconds  
(without alignments)  
9900.026 Million cell updates/sec

Title: US-09-913-767-10

Perfect score: 1067

Sequence: 1 agacaagaatgtgaaggct.....gaccataagtgaagttaa 1067

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3413475 seqs, 253800928 residues

Total number of hits satisfying chosen parameters: 6826950

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications\_NA.\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq.\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq.\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq.\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*
- 13: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq.\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq.\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq.\*
- 16: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq.\*
- 17: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq.\*
- 18: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*
- 19: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*
- 20: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*
- 21: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	317.8	29.8	1233	16	US-10-424-599-30274
2	284.8	26.7	1146	16	US-10-425-114-10516
3	248.6	23.3	994	16	US-10-424-599-73501
4	194.4	17.2	1173	17	US-10-437-963-57539
5	182.2	17.1	1513	17	US-10-437-963-82061
6	161.8	15.2	2214	16	US-10-424-599-102519
7	145.6	13.6	3387	9	US-09-938-842A-2212
8	145.6	13.6	3387	11	US-09-938-842A-2212
9	136.4	12.8	1373	16	US-10-424-599-31364
10	133.6	12.5	1777	17	US-10-437-963-91425
11	133.6	12.5	1417	16	US-10-424-599-97043
12	132.8	12.4	789	17	US-10-437-963-72452

13	123	11.5	1296	16	US-10-425-114-13687	Sequence 13687, A
14	120.4	11.3	1333	16	US-10-425-114-10574	Sequence 10574, A
15	119.4	11.2	737	16	US-10-425-114-24295	Sequence 24295, A
16	114	10.7	83698	17	US-10-416-898-9	Sequence 9, Appli
17	112.2	10.5	716	16	US-10-424-599-61310	Sequence 61310, A
18	111	10.4	2928	17	US-10-437-963-33800	Sequence 33800, A
19	110	10.3	1355	16	US-10-424-599-31363	Sequence 31363, A
20	104.6	9.8	1294	16	US-10-424-599-2524	Sequence 2524, Ap
21	103.8	9.7	997	16	US-10-424-599-9213	Sequence 75161, A
22	98	9.2	792	16	US-10-424-599-9213	Sequence 9213, Ap
23	92.8	8.7	2381	17	US-10-437-963-43860	Sequence 43860, A
24	84.2	7.9	572	16	US-10-424-599-118451	Sequence 118451, A
25	77.2	7.2	1306	17	US-10-437-963-18858	Sequence 18858, A
26	74	6.9	747	17	US-10-767-701-6806	Sequence 6806, Ap
27	72	6.7	592	9	US-09-770-152-44	Sequence 44, Appl
28	67.2	6.3	522	16	US-10-424-599-114931	Sequence 114931, A
29	56.2	5.3	1119	17	US-10-437-963-34510	Sequence 34510, A
30	55.2	5.2	1386	16	US-10-425-114-12485	Sequence 12485, A
31	54.2	5.1	1122	17	US-10-437-963-42424	Sequence 42424, A
32	52.4	4.9	728	17	US-10-767-701-9000	Sequence 9000, Ap
33	52	4.9	1779	16	US-10-425-114-28473	Sequence 28473, A
34	51.6	4.8	2605	17	US-10-437-963-83900	Sequence 83900, A
35	51.4	4.8	486	16	US-10-424-599-74177	Sequence 74177, A
36	47	4.4	774	9	US-09-770-445-895	Sequence 895, App
37	46	4.3	327	16	US-10-424-599-121152	Sequence 121152, A
38	42.2	4.0	41104	9	US-09-816-685-3	Sequence 3, Appli
39	42.2	4.0	41104	16	US-10-639-708-3	Sequence 3, Appli
40	41	3.8	1432	17	US-10-437-963-1897	Sequence 1897, Ap
41	40.6	3.8	6317	15	US-10-204-708-11	Sequence 11, Appl
42	40.6	3.8	6317	15	US-10-311-455-381	Sequence 381, App
43	40.6	3.8	94512	13	US-10-087-192-976	Sequence 976, App
44	39.2	3.7	478	10	US-09-918-995-22681	Sequence 22681, A
45	39	3.7	470	13	US-10-027-632-248589	Sequence 248589, A

ALIGNMENTS

RESULT 1  
US-10-424-599-30274  
; Sequence 30274, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 30274  
; LENGTH: 1233  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_12733C.1  
; US-10-424-599-30274

Query Match	29.8%	Score 317.8;	DB 16;	Length 1233;
Best Local Similarity	61.3%	Pred No. 1.8e-83;		
Matches	568;	Conservative	0;	Mismatches 347;
			Indels	12;
			Gaps	3;
Qy	79	TGCGTCTCTACTTCAACAAATCGCGTAAAGAGATTGGTTCTCTACGTTTCTTGAACCTG	138	
Db	2	TGCGTCTCTACTTCAATCCACGGTGCAGAAACGAATCTGCTCTCTAGCTTCTTGAACCTG	61	
Qy	139	CAGGCTTCTCTGTTATCTTCAATCTCTGCTCTCTCTTCTTACATACCGCGGCAGAACCA	198	
Db	62	CAGCAATTCGGTCGAGTCTCTCCCAATCAATCAATTTCTTACATT-----CATAAACG	113	





Db 588 CGCGGGACCGCGCGGGGTCTCGCGCGCAGTACTCGCGGGTTCGCCATGAC 647  
Qy 530 TGTAGACAGAGTGTTATGTTTTCATCTTGCCATTAGTGGAATTGCTTACCAGAA 589  
Db 648 GCTCGCGCGCGCGGTGTACGGGCTCGTGTCTCCCGTCTATGGAGCTTCAGCCAGGGCA 707  
Qy 590 -----ACTAAGCAAAACATGAGCTATACCTTGTGTCTCGAGTTCACAGTTGATTTT 640  
Db 708 CCACGCCCGCGCGCGCGGTACAGTACACGCTCGTCAATGAGATGCGAGCTCGTCAAT 767  
Qy 641 GTGTCTCTCTGCTTCTATGTGTCAGCGTCAATCGGTATGTTTCAATCGTGTGATTTCAAGCA 700  
Db 768 CGGCTTCGTGCCAGCGCTTCAGCGCGTTCGGCATGCTCGTCAACACGATTTTCA 824  
Qy 701 GGCCTTACAAAAGCAAGAGAGTTCAGCTTGGAGAGCATGTTTCTATGTTGGTGGC 760  
Db 825 CGCAATCCAGAGAAGCCATGAATTCGGTCTCGGGCCAGCGGCTACTACCTGCTCCT 884  
Qy 761 TGTGTTTTCCAGCATCATATGGCAAGGCTTCTTCTGGAGCCATTGGATTAATCTTCTC 820  
Db 885 GGCGGATCGCGCGCGCTGTACCAAGTGTCTTCTCGGCACGATCGCGCGCATCTTCTA 944  
Qy 821 CACATCGTCTCTCGTCTCGGTATATGATATCAAGTGTCTTTCGCAATTACAGAGGTTT 880  
Db 945 CGGCTCGCGCTGCTCGCGCGCTCATATGACCGTGTCTATCCCGTTCACCGAGTGTCT 1004  
Qy 881 AGCTGTTATATCTACCAATGAAAAGTTTCAAGCTGAGAGGAGCATTTCTCTGTCTCTC 940  
Db 1005 CGCGCTCATGTTCTTCCACGAGCCATTCAACGGCACAAAGGCGCTCGCGCTCTCTC 1064  
Qy 941 CCTTTGGGCTTGTCTCTTCTTCTTATGTTGAGATAAAGTCTGCGGAGG 990  
Db 1065 GCTATGGGCTTCGTCTCTTCTTCTTCTTCTACGGCGAGTTCGGGCGGCAAGG 1114

RESULT 5  
US-10-437-963-82061/c  
; Sequence 82061, Application US/10437963  
; Publication No. US2004012343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Boukharov, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 82061  
; LENGTH: 1513  
; TYPE: DNA  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_81525C.1  
US-10-437-963-82061

Query Match 17.1%; Score 182.2; DB 17; Length 1513;  
Best Local Similarity 50.9%; Pred. No. 4.8e-43;  
Matches 458; Conservative 0; Mismatches 438; Indels 3; Gaps 1;  
Qy 63 GGAGTCTCTTGGTATGCTCTACTTCAACATGCGCGGTAAAGGATTTGCTCTCT 122  
Db 1207 GCATCTACATCTTGGGAGGTCTTCTACATCAAGCGCGCAATAGCAAGTGTGTC 1148  
Qy 123 ACGTTCTTTGAACTGCAAGGCTTCTCTGTTATCTTCAATCTCTGCTCTTCTTACATT 182  
Db 1147 ACAITCGTCCAACTGCTGCGCTTCCGATTTTGTTCATTGCGCTTCTTCTTCCATTCA 1088

Qy 183 ACCGGCGCAGAACCAATATGTTGGGTGATAGTACAAGTTTCTTTCTTATCAAAACCGGT 242  
Db 1087 AAGCATCTTTCTACACAAACAGTCACTAGTAGTCTGCTCCCTCAAAATTTCTATCCCAA 1028  
Qy 243 CTTCCTATCGCGGTGTTATGTTAGGCATCTCTCAGGTTTGTATTAATCTTCTGATGCA 302  
Db 1027 ATTACTCTGATATATGTTTCTCGGCTCATCATTTGCTCGACAGCATTTGATGATTC 968  
Qy 303 TATGTTATAGCTTATCTTCCAGTTTCTACAGCTGCTCTTATCATTTGCTTCAAGTTAGCT 362  
Db 967 TATGGCTACTATATCTTCGGTCTCAACATATTCGCTCATCTGTGCTAGTCAGCTTGC 908  
Qy 363 TTTATAGCTATCTTCTCATTTCTTCAAGTTTAAACATAAAGTTTCACTCTTTTACCATCAAT 422  
Db 907 TTTCAATGCTCTCTCTCATATTTCTCAATGCTCAAAATTCACCCCTCTGATTTTCAAT 848  
Qy 423 GCTGTTGTTGTTGACTGTTGCTGCGGTTTGGGAATGATACCGAACTGATAAG 482  
Db 847 TCCGATGCTCTCTTACGTTTCTGCTTCACTCTTGGAGTTGATGAAGATTCTCAGGGA 788  
Qy 483 CCAGTTTCATGAGACTCAACAGCAGTACATAACTGGTTTCTTGTATTTCTGTAGCAGCT 542  
Db 787 ACTACTAGTATATACATGGAAGTACATTTTGGGTTTCTGTTGACACTAGGGCATCA 728  
Qy 543 GTTATGATGCTTTCATCTTTCGCTTGTAGTGAACCTTACAGAAAGTAAAGCAAAACC 602  
Db 727 GCTACATACCTCGCTCATTTCTCTCCGTGATGCAAGTCAATTTGAGAAGGTTATTAAGAG 668  
Qy 603 ATGAGCTATACCTTGTGTCGAGTTCAGTTGATTTTGTGTTCTCTTCTTATTTGTC 662  
Db 667 GAGACCTTCTCAGTTGTTGAACTGACATATATACAGCTCTCGTGGCAACATTGGCT 608  
Qy 663 AGCCTCATCGGTATGTTTCATCGCTGCTGATTTTCAAGCAGCCCTTACAAAAGAAAGCAAGA 722  
Db 607 TCTCTTGTGGTTATTTGCAAGTGGTGAATGATGA ---CTTTACAAGAGAGATGCAT 551  
Qy 723 GAGTTCAAGCTTGGAGGCAATTTCTATGTTGGTGGCTGTTTTCAGCAATCATATGG 782  
Db 550 GCATTTCCAGTCTGGAAAGCTGTATATGTAATGACACTGCTGTGGAGCGGTATATCTTGG 491  
Qy 783 CAAGGCTTCTTCTTGGAGCCATTGGATTAATCTTCCACATGCTCTCTGCTCGGT 842  
Db 490 CAGGTAGCATCAGTTGGAGTGGGATGATCTTTGTGGTGTGATCGCTGTTTCAAT 431  
Qy 843 ATTATGATATCAGTGTCTTTTGCCTTACAGAGTTTTCAGTGTATATTTCTACCATGAA 902  
Db 430 GTGATAAGCACCTTAGCTCTACCCATCATTTCTGTTTGTGCTGTGATTTTCTTTCATGAC 371  
Qy 903 AGTTTCAAGCTGAGAGGCACTTTCTCTGCTCTCTCCCTTTGGGCTTTGCTCTTA 961  
Db 370 AAGATGGATGGAGTAAAGATTATAGCTATGCTGATGCCAATTTGGGGAATTTATGTCTATA 312

RESULT 6  
US-10-424-599-102519  
; Sequence 102519, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 102519  
; LENGTH: 2214  
; TYPE: DNA  
; ORGANISM: Glycine max



```

; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_63592C.1
US-10-424-599-102519

```

Query Match	15.2%	Score 161.8	DB 16	Length 2214
Best Local Similarity	51.3%	Pred. No. 7.1e-37		
Matches 487	Conservative 0	Mismatches 442	Indels 20	Gaps 4
QY	20	TCTTGTGATCATAAACTGCATAAATTCCTAGCCATAGGAAACTGTGAGGTCTCTTGATTAT	79	
DB	416	TCCTGTGACACTCAGCATAGCCTTCTTATAGTTGGCCATCTGCTGCTGTTATCCTTGG	475	
QY	80	GGGCTCTACTTCAACAATGCGGTAAAGGATTTGGTTCTCTACGTTTCTTGAACATGC	139	
DB	476	AAGATTTTATATGATCAGGGTGAATAAGTAAATGGATGCTACTCTAGTTCAAACTGC	535	
QY	140	AGGCTTTCCTGTTATCTTCAATTCCTCTGCTCTTCTCTTACATTAACCGGCGCAGAGCA	199	
DB	536	TGCTTCCCGACTTGTTCATTTCCATTATT-TACAAATTCCTTACCCTCCAGAGGCTTCA	594	
QY	200	CAATGTGGGTGATAGTACAAAGTTTCTTTCTTATCAAAACCGGCTCTTCTATCGCGGTGT	259	
DB	595	C-----TTCTGCTTCACCTCCCATCAAAATATTCCTTTGATATATTTGG	640	
QY	260	TATTTGAGGCATTTCTCAGGGTTTGATAACTACTTGTATGCAATAGGTATAGCTTATCT	319	
DB	641	TCT--TTGGAGTCTTAAATGCTGCTGACAATATGATGACTTCCACTGGACTCTTATACCT	698	
QY	320	TCCAGTTTCTACAGCTGCTCTTATCATTTGCTTACAGTTACAGTTAGCTTTTATAGCTATCTCTC	379	
DB	699	CTCGCTTCTACCTTAITTCGCTGATTTGTGCAATCACAGTTAGCTTTTAAATGCAGTTTCTC	758	
QY	380	ATTCTTTCATGGTTAAACATAAAGTTTCACTCTCTTTTACCATCAATGCTGTGTGTTGTTGAC	439	
DB	759	ATATTTTATCAAATCTCAAAGTTTCACTGCCTTGATTTATAAACTCTACAGTGGTTCTCAC	818	
QY	440	TGTTGTGCTGCGGTTTTTGGGAATGCAATACGAAACTGATAAGCCAGTTTCATGAGACTCA	499	
DB	819	TTTATCTGCTGCACTCTCTGCTGTTAAACGAAGACACAGATGAACCATCTGGTTTCTCCAA	878	
QY	500	CAAGCAGTACATAACTCGTTTCTTGATTTACTGTAGCAGCAGCTCTGATGTATGCTTTTCAT	559	
DB	879	GGGAAGTACATATTTGGTTTCTCTATGTACCCCTTGGAGCTTCTGCAGTGTACTCTCTTTT	938	
QY	560	CTTGCCATTTAGTGGAACTTGCTTACCAGAAAGCTTAAGCAAAACCATGAGCTATACCTTGT	619	
DB	939	GCTTTCCCTCATGCAGCTGACCTTTGAGAAGTTCTGAAGAAGAAACATTTCTGTGTTG	998	
QY	620	GCTCGAGTTCAGTTGATTTTGTGCTCTCTTGCTTCTATTGTCAAGCTCATCGGTATGTT	679	
DB	999	TTTGGAAATGCAAACTTACACATCATTTCTGTTGCCTCTGGTCTTCTGTCTATPAGGCCATT	1058	
QY	680	CATCGCTGGTGTATTTCAAGCAGGCTTTACCAAAAGAAAGCAAGAGTTTCAAGCTTGGAGA	739	
DB	1059	TGCAGTGGGAATG---GGGATTTTGCATGCGAATAAGGAGGTTTTCAGAAAGATA	1115	
QY	740	GGCATTTGTTCTATGTGTGTGGCTGTGTTTTAGCCATCATATGGCAAGGCTTCTTCTTGGG	799	
DB	1116	TGTTGCTTATGTTATGACTTTTGGTTTGGACTTCAATAGCCTGGCAGGTATGCTCTGTGG	1175	
QY	800	AGCCATTTGGAATTAATCTTCCACATCGTCTCTCGTCTCGGGTATTTATGATACAGTGCT	859	
DB	1176	TGTTGTTGGCTTGATCTTCTTAGTGTCTTCTCTACTTCCAAATGTTTAAAGCACAGTTTC	1235	
QY	860	TTTGCCAAATACAGAGGTTTTAGTGTGTATATTTCTACCATGAAAAGTTTCAAAGCTGAGAA	919	
DB	1236	TTTAGCCGTAACCTCTATGCTGCTGTTATAGTTTTCATGATGAATGAATGGGTGAA	1295	
QY	920	GGGACTTTTCTGCTCTCTCCCTTTGGGGCTTTGCTCTCTTACTTTTAT	968	
DB	1296	GATAAATTTATGCTTTTGGCTCTATGGGGTTTTGGCTCTTATATTTAT	1344	

```

RESULT 7
US-09-938-842A-2212
; Sequence 2212, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kieps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIPT300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 2212
; LENGTH: 3387
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-2212

```

Query Match	13.6%	Score 145.6	DB 9	Length 3387	
Best Local Similarity	49.1%	Pred. No. 6.1e-32			
Matches	475	Conservative 0	Mismatches 484	Indels 9	Gaps 3
Qy	1	AGACAGAATGGTGAAGGCTCTTGATCATAACTGCATAATCTAGCCATAGGAACT	60		
Db	2354	AGAACTGTAAGAGGTGGCTCGGTCTCCATATACGCAATCTTTGTGCATCTTCTGCCAAC	2413		
Qy	61	GTGAGATCCTTTGATTAATATCGCTCTCTACTTCAACAATGGCGGTAAAGGATTTGGTTCT	120		
Db	2414	CACCTGCTACAGTTCTGGGTAGACTGTACTATGAAATGAGGGAAAGCACATATGTGG	2473		
Qy	121	CTACGTTTCTTGAACCTGACAGGCTTCCCTGTTATCTTCACTCTCTGCTCTTCTCTTACA	180		
Db	2474	TAAACACTTCTTCAACTCATTTGGCTCCCTGTACTGATCTCTTCCGCTTCTTTCTCGAA	2533		
Qy	181	TTACCCGGCGCAGACGAACAATGGGGTATAGTACAAAGTTTCTTCTTATCAAAACCGC	240		
Db	2534	TCAGGCAACCCAATCAACAGATACAAATTTTCAGT--CAGTCCCTTCTCTCACCACTCT	2591		
Qy	241	GTCTTCTTATCGCGCGTGTATTGTAGGCAATCTCTCAGGGTTTGTAACTACTTGTATG	300		
Db	2592	TGCATCGTTTACTTGTGCACT---GGACTGCTAGTGTCTGCTTATGCTTATTTGCTG	2647		
Qy	301	CATATGGTATAGCTTATCTTCCAGTTTCTACAGCTGCTCTTATCATTTGCTTTCACAGTTAG	360		
Db	2648	CAGTTGGTTTGGTTTACTTTACCACTCTACTTTTCCCTCATCTTGGCCTCACAGTTGG	2707		
Qy	361	CTTTTATAGCTATCTTCTCAATCTTCATGGTTAAACATAAAGTTTCACTTCTTTTACCATCA	420		
Db	2708	CCTTCACTGCTTTTCTCATATTTCTTAACTCGGAAAGTTTCACTTCTTTGATGTCA	2767		
Qy	421	ATGCTGTGTGTGTTGTGACTGTTGGTGCTGCGGTTTTGGGAATGCATACCGAAACTGATA	480		
Db	2768	ATTCTTTGTTTCTCCTTACGGTTTCTCTGSCCTCTCTCGTGGTCAACACTGATTTCAGAA	2827		
Qy	481	AGCAGTTCAATGAGACTCACACAGTACATACTACTGGTTTCTTGATTTACTGTAGCAGCAG	540		
Db	2828	ACACAAACAATGTATCTAGAGTACAAATATGTGATTTGGGTTTCATCTGTACCAATGGTGCTT	2887		
Qy	541	CTGTATGTATGCTTTTCATCTTCCCATTAGTGGAACTTGCTTACCAGAAGCTTAAGCAAA	600		
Db	2888	CCGCTGGGATTGGATTGGTACTATCTCTGATACAACTGCTCTTTCAGAAAGTTTTACGA	2947		
Qy	601	CCATGAGCTATACCCCTTGCTCGAGTTCCAGTTGATTTTGTGCTCTCTTGTCTTATTG	660		

Db 2948 AGCATACATCTCAGCAGTCTCGGACTTGGCCAAATACCAAGTCTCTAGTTCACAACTGTG 3007  
Qy 661 TCAGGCTCATCGGTATGTTTCATCGCTGCTGATTTCAACAGCGCCCTTACCAAAAGAGCAA 720  
Db 3008 TGGTACTCATAGGACTGTGTTCAAGTGGAGAGT---GGAGAACTCTGCAAGTGAGATGA 3064  
Qy 721 GAGAGTTCAAAGCTTCGAGAGGCAATGTTCTATGTGCTGCTGTTTTCAGCCATCATAT 780  
Db 3065 GAACTACAAGTGGGAAAGTGTCATATATCTTGACTTTGGCCCTCAGCAGCTATTTTCT 3124  
Qy 781 GCGAAGGCTTCTCTTGGAGCCATGGAATTAATCTCTCCACATCGTCTCTCGTCTCGG 840  
Db 3125 GCGAAGTATACACTGTTGTTGTGGGATTAATCTTCAGAGTCTCTCTGTGTTCCCA 3184  
Qy 841 GTATTATGATATCAGTGTCTTTGGCAATTCAGAGGTTTTCAGCTGTGTATATTCACCATG 900  
Db 3185 ATTCCATAAATGCTGTGGGACTGCTTATAGTTTCAGAGTGTAGCAGTGTAGTTTCCATG 3244  
Qy 901 AAAAGTTTCAAGCTGAGAGGAGCTTCTCTTGTCTCTCTCCCTTTGGGGCTTTGCTCTT 960  
Db 3245 ATAAGATGATGATCAACAAAGATTTTCTCCATCATTTTAGCTATCTGGGGCTTCCCTTCA 3304  
Qy 961 ACTTTTAT 968  
Db 3305 TTGTCTAT 3312

## RESULT 8

US-09-938-842A-2212  
; Sequence 2212, Application US/09938842A  
; Publication No. US20040009476A9

## GENERAL INFORMATION:

; APPLICANT: Harper, Jeff

; APPLICANT: Kreps, Joel

; APPLICANT: Wang, Xun

; APPLICANT: Zhu, Tong

; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING

; FILE OF INVENTION: SAME, AND METHODS OF USE

; CURRENT APPLICATION NUMBER: US/09/938,842A

; CURRENT FILING DATE: 2001-08-24

; PRIOR APPLICATION NUMBER: US 60/227,866

; PRIOR FILING DATE: 2000-08-24

; PRIOR APPLICATION NUMBER: US 60/264,647

; PRIOR FILING DATE: 2001-01-16

; PRIOR APPLICATION NUMBER: US 60/300,111

; PRIOR FILING DATE: 2001-06-22

; NUMBER OF SEQ ID NOS: 5379

; SEQ ID NO 2212

; LENGTH: 3387

; TYPE: DNA

; ORGANISM: Arabidopsis thaliana

US-09-938-842A-2212

Query Match 13.6%; Score 145.6; DB 11; Length 3387;

Best Local Similarity 49.1%; Pred No. 6.1e-32;

Matches 475; Conservative 0; Mismatches 48; Indels 9; Gaps 3;

Qy 1 AGACAAGAAATGGTGAAGGCTCTTGTGATCATAAACTGCATAAATCTAGCCATAGGAAACT 60  
Db 2354 AGAACTGTAAGAGGTGGCTCGGTGCTCCATATAGGCAATCTTGTCTATCTTCTGCCAAC 2413  
Qy 61 GTGGAGGTCCCTTGNATTAATGCTCTACTTCAACAAATGGCGTAAAGAGTATGGTTCT 120  
Db 2414 CACTTGTACAGTTCTCGGTAGACTGTACTATGAAATGGAGGAAAGACATATGTGG 2473  
Qy 121 CTACGTTTCTTGAAGCTTCAGGCTTCTGTTATCTTCAATTCCTCTGCTCTTCTTCTACA 180  
Db 2474 TAACACTTCTTCAACTCAATTCGCTTCCCTGCTACTGATTTCTCTCGGCTCTTCTCGAA 2533  
Qy 181 TTACCCGCGCAGAGCAACAAATGTTGGGTGATAGTACAAAGTTCTTTCTTATCAAAACCGC 240  
Db 2534 TCAGGCAACCCAAATCAACAGATACAAATTTTCAGT--CAGTCCCCTTCTCTTCCACCCT 2591





```
QY 261 ATTGTAGGATCTCTCAGGGTTGTATACTATTGTATGCATATAGTATAGCTTATCTT 320
Db 451 GTGCTCGGCTCTCTCAACGGCGCGACGACTTGGTCTACGCTACGGCTCGGCTACCTG 510
QY 321 CCAGTTTCTACAGCTCTCTATCATTTGCTTCAAGTTTCAAGTTTATAGCTATCTTCTCA 380
Db 511 CCGGTGTCACCTCCGCGCATCTCTCATCTCACGAGCTCGGTTTCAACCGTCTCTTCGGG 570
QY 381 TTCTTCACTGTTAACTAAAGTTTCACTCTCTTTTACCATTCAATCTGTTGTGTTGACT 440
Db 571 TGCTCATCTGCGGAGCGGCTCACCGGCGACGCTGAACGCGGTGGCGTCTCACT 630
QY 441 GTTGTGCTCGGTTTGGGAATGCTATACCGAACTGATAAGCAGTTTATAGAGACTCAC 500
Db 631 ATCGGCGCTCTGCTGGGCTTCCACGCGTCCAAGGACCGACCCGCGGGTGACCAAC 690
QY 501 AACGAGTACATACTGGTTTCTTGATTACTGTAGTACGAGCTGTTATGTATGCTTTCATC 560
Db 691 GGAAGTACTGGATGGGCTTCTTCTCACCTCGGCGCGCGGCTGTACGGGCTCATC 750
QY 561 TTGCCATTAGTGAAGTTGCTTACCAGAAAGC 592
Db 751 CTGCGCTCGTCGAGCTCGGTACAGCAGC 782

RESULT 13
US-10-425-114-13687
; Sequence 13687, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 13687
; LENGTH: 1296
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB143-024-D8_FLI
US-10-425-114-13687

Query Match 11.5%; Score 123; DB 16; Length 1296;
Best Local Similarity 50.9%; Pred. No. 1.8e-25;
Matches 350; Conservative 0; Mismatches 325; Indels 12; Gaps 2;

QY 21 CTGTGATCATAAACTGCATAATTTCTAGCCATAGGAACTGTGGAGTCTTTTGATTATG 80
Db 88 CTGTGGCGACCAACTGGGTGATGCTCGGCTTGGCGTCAACCGTGGGAGCTCTCTCAGC 147
QY 81 CGTCTACTTTCAAGATGGCGGTAAAGATTGGTTCTCTACAGTTTCTTGAACTGCA 140
Db 148 CGCTCTACTTTTAGCAAGGCGGACACCGGCACTGGCTCTCCGGATGGGTCTCCAGACCGGT 207
QY 141 GGTCTTCTGTTATCTTCACTCTGCTCTTCTTATCAATTAACCGGCGGAGCAAC 200
Db 208 GGTGGCGGCTGCTGCTCCGCGAGTGGGCGTCTCTACGT---CCGCGCGGAGCGCGC 264
QY 201 AATGTGGGTGATAGTACAAGTTTCTTTCTTATCAAAACCGGCTCTTCTTATCGCGCTGT 260
Db 265 GACCGAGGCGCCCGGTGCTCACCCAGACCCGCGGATCTGCTGGCGCGCG 324
QY 261 ATTGTAGGATCTCTCAGGGTTGTATACTATTGTATGCATATAGTATAGCTTATCTT 320
Db 325 GGGCTCGGGCTCATCGGCGCGTGGACAACTTGTCTACGCTTGGGCGCTCGAGTTCTCTG 384
```

```
QY 321 CCAGTTTCTACAGCTGCTTATCATTTGCTTCAAGTTTATAGCTTATAGCTATCTTCTCA 380
Db 385 CCGGTCTCAACCTCCGCGATCTCATCTCCACGAGCTGGCCTTCAAGTGTCTTTCGGG 444
QY 381 TTCTTCACTGTTAACTAAAGTTTCACTCTCTTTTACCATTCAATGCTGTTGTGTTGACT 440
Db 445 TACCTGATGCTGCGGAGCGGCTGACGATGGGCGGTAACCGCTGGCGTCTGCTGAGC 504
QY 441 GTTGTGCTCGGTTTGGGAATGCATACCGAACTGATAAGCAGTTTCAAGAGACTCAC 500
Db 505 GTGGGCGGCTGCTGCTGGGCTGATGTCTCTCTGACCGCGCGCGGCGTCAACGAG 564
QY 501 AAGCAGTACATAAAGTTTCTTGTATCTGTAGCAGCAGCTGTTATGATGCTTTCATC 560
Db 565 AGTCAGTACTGGCTGGCTTCAAGCTCACCTGGCGCGCGGCTGTGTACGGGCTGTTT 624
QY 561 TTGCCATTAGTGAAGTTGCTTACCAGAAAGCTAAG-----CAAACTATGAGCTAT 611
Db 625 CTGCGGCTGGTCAAGTCACTTACCAAGTGGCGGCGGCGGCGGAGCGCCCGTGAAGTAC 684
QY 612 ACCCTTGTGCTGAGTTCCAGTTGATTTTGTGCTCTCTTGTCTTATGTCAGCGCTATC 671
Db 685 GCGTAGTGTGGAGTTGCAGCTGGTGTGAGTGGGTTTGTGCGCACCGCGTTTCTGACCGCTC 744
QY 672 GGTATGTTTCACTGCTGCTGCTGATTTCAG 698
Db 745 GCGATGATCGTCAACAGGACTTCCAG 771

RESULT 14
US-10-425-114-10574
; Sequence 10574, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 10574
; LENGTH: 1333
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700944981_FLI
US-10-425-114-10574

Query Match 11.3%; Score 120.4; DB 16; Length 1333;
Best Local Similarity 49.0%; Pred. No. 1.1e-24;
Matches 350; Conservative 0; Mismatches 361; Indels 3; Gaps 1;

QY 257 TGTATTGTAGGATCTCTCAGGGTTGTGATAAATCTTGTATGCATATAGTATAGCTT 316
Db 462 TATCGTTTGGGTTTCTTAAGTGTCTGATAAACCTCATGTATGTATGCTTATGCTATGCTTA 521
QY 317 TCTTCCAGTTTCTACAGCTGCTTATCATTTGCTTTCACAGTTAGCTTTTATAGCTATCTT 376
Db 522 CCTCCCTGCATCACTGCTCTCACTTGTGGCATCATCCCTTGTGTTTTCGCGCTCTT 581
QY 377 CTCATTTCTATGGTTAACTAAAGTTTCACTCTCTTTTACCATCAATGCTGTTGTGTTGT 436
Db 582 TGGATACTTTCTTGTGAAGAACAAAGTGAATGCTTCAATAGTAATTCGTTTTCGTAT 641
QY 437 GACTGTTGGTGTGCGGGTTTGGGAATGCATACCGAACTGATAAGCCAGTTTCATGAGAC 496
```

```
Db 642 AACCGTGCAATGACCAATCAATGCGCTGGACTCGAGTTCAGACAGATATCCAGCAATTAG 701
Qy 497 TCACAGCAGTACATAAATCGGTTTCTTGATTACTGTAGCAGCAGCTGTTATGTATGCTTTT 556
Db 702 TGACAGTGAATACATCATGGGATTTGTATGGATGTTTATAGGATCTGCTTCCATGGGCT 761
Qy 557 CAATCTGGCAATTAGTGGAACTTGTTCACAGAAAGCTAAGCAAAACCATGAGCTATACCCT 616
Db 762 TATTTTCGCTCTCTCAGAGCTCGTCTTGTGAAGTGTCTCGAAGAAGATCCTTTATCGT 821
Qy 617 TGTGCTCCAGTTCAGATTTTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 676
Db 822 TGTCTGGAGCAGCAAGTCATGGTTCTCTATTTGCAATTTCTGTTTACCACCTGTAGGGAT 881
Qy 677 GTTCATCGTGTGTGATTTCAAGCAGGCTTACCAAAAGCAAGAGAGTTCAAGCTGG 736
Db 882 GATTGTGAGTGTGATTT---TCAGGAGATGCCACATGAGGCTACCATTTTCGAAAGTGG 938
Qy 737 AGAGCAATGTTCTATGTGGTGGCTGTGTTTTCAGCCATCATATGGAAGGCTTCTTCTT 796
Db 939 TAGAAGTGTCTTATATCTGTTATCATTTGGGGTGCATCACTTTTCAGCTGGGGTTCT 998
Qy 797 GGGAGCCATTGGAATTAATCTTCTCACATCGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 856
Db 999 GGGGGGCACTGTATAAATTTTCTGGGCTCTACTGTGTAGCAGGTGTGCTTAATGCACT 1058
Qy 857 GCTTTTGGCAATTACAGAGGTTTTCAGTGTGTATATTTACCATGAAAGTTTCAAGCTGA 916
Db 1059 AAGAACCCATTAACAGCATTTGACGCTGTTATCTGCTAAAGGACCCTATGAGTGGTTT 1118
Qy 917 GAAGGCACTTCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 970
Db 1119 CAAGATCCTCTCCCTAGTGATCACCTTTTGGGATTTGGCTCATATATTTATGG 1172
```

RESULT 15  
US-10-425-114-24295  
; Sequence 24295, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 24295  
; LENGTH: 737  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB3608-022-E3\_FLI  
US-10-425-114-24295

Query Match 11.2%; Score 119.4; DB 16; Length 737;  
Best Local Similarity 52.0%; Pred. No. 1.5e-24;  
Matches 298; Conservative 0; Mismatches 266; Indels 9; Gaps 1;  
Qy 20 TCTTGTGATCATAAATCTGATTAATCTAGCCATAGGAACTGTGGAGTCTCTTTGATTAT 79  
Db 165 TCTGTAGCGTCAACTCGGGATGCTGCGGTGGAGCCATCGCGGGCGGCTCCTGAG 224  
Qy 80 GCGTCTCTACTTCAACAATGGCGGTAAAGGATTTGTTTCTCTAGCTTTCTTTGAAACTGC 139  
Db 225 CCGGCTCTACTTACAGAGGGGGGACCGGAGTGGCTGTCCGGGTGGTGGAGACGG 284  
Qy 140 AGGCTTTCCTGTTATCTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 199

```
Db 285 CGGGTGGCGCTGCTGCTGATTCCCGTGGCGGCTGCTTCTGTGCGCGCGCGCGCGA 344
Qy 200 CAATGTGGGTGATAGTACAAAGTTTCTTTCTTATCAAAACCGGCTCTTCTTATCGCGGCTGT 259
Db 345 CAGGG-----GCGCCCCGGTCTCTGCTGCGCGCGCGCGGATCTGCTGGCGCGCGC 395
Qy 260 TATTGTAGGCATTTCTCTCAGGGTTTGTAACTACTTGTATGCATATGATATAGTATTATCT 319
Db 396 GGGCTCGGCTGGCCACGCGGCGGAGACTTCACTACGCCCTTCGGGCTCGGCTACCT 455
Qy 320 TCCAGTTTCTACAGTCTCTTATCATTTGTCTTCAAGTACGTTTATAGCTATCTTCTTC 379
Db 456 CCGGTGTCCACGTCCGCGATCCTGATCTCGACGAGTGGCGTTACGGGTGTTCTTCGC 515
Qy 380 ATTCTTTCATGGTTAAACATAAAGTTTCACTCTCTTTTACCATCAATGCTGTTGTGTTGAC 439
Db 516 GTTCTCTGTTGGTGGCGAGCGCTGACGCGCGCTCCGTGAAACCGCGCTGCTGCTGAC 575
Qy 440 TGTGGTGTCTGCGGTTTTGGGAATGCATACCGAACTGATAAGCCAGTTTCATGAGACTCA 499
Db 576 GGTGGGCGCGTGGTGTCTGGGCTGCGAGTCTCGTCGACCGGCCCGGCGGCTCACAG 635
Qy 500 CAAGCAGTACATAAATCTGTTTCTTGTATTTACTGTAGCAGCAGCTGTTATGATGCTTTCA 559
Db 636 GGGCCAGTACTGGCTGGGCTTCTGCTGACGCTCGGCGCGCGCGCTCTACGGCCTCGT 695
Qy 560 CTTGCCATTAGTGAACCTTCTTACACAGAAAGC 592
Db 696 CTTGCCCTCTGTGGAGCTCGCTACAGGCGGC 728
```

Search completed: November 1, 2004, 23:40:13  
Job time : 555.64 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 1, 2004, 05:48:03 ; Search time 98.4346 Seconds  
(without alignments)  
7704.726 Million cell updates/sec

Title: US-09-913-767-10  
Perfect score: 1067  
Sequence: 1 agacaagaatggtgaaggct.....gaccataagtgaagttaa 1067

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA.\*  
1: /cgn2\_6/ptodata/1/ina/5A-COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B-COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A-COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B-COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PCTUS-COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	59.4	5.6	7218	1 US-08-232-463-14	Sequence 14, Appl
2	43	4.0	1141	4 US-09-806-708B-22	Sequence 22, Appl
C 3	41.6	3.9	289	3 US-09-007-005-17	Sequence 17, Appl
C 4	41.6	3.9	289	3 US-09-244-796-17	Sequence 17, Appl
5	40.6	3.8	6317	4 US-10-204-708-11	Sequence 11, Appl
C 6	37.4	3.5	933	4 US-09-248-796A-13724	Sequence 13724, A
7	37.2	3.5	1055	4 US-09-806-708B-23	Sequence 23, Appl
C 8	37.2	3.5	1664976	4 US-08-916-421B-1	Sequence 1, Appl
C 9	37.2	3.5	1664976	4 US-09-692-570-1	Sequence 1, Appl
10	37	3.5	1308	4 US-09-270-767-12733	Sequence 12733, A
C 11	37	3.5	4182	3 US-07-973-257-1	Sequence 1, Appl
C 12	36.2	3.4	1738	3 US-09-202-712-18	Sequence 18, Appl
C 13	35.8	3.4	399	4 US-09-621-976-8976	Sequence 8976, Ap
C 14	35.8	3.4	832	4 US-09-621-976-2813	Sequence 2813, Ap
C 15	35.4	3.3	534	4 US-09-248-796A-751	Sequence 751, Ap
C 16	35.4	3.3	832	4 US-09-621-976-2813	Sequence 2813, Ap
C 17	35.2	3.3	1485	3 US-09-134-001C-1638	Sequence 1638, Ap
18	35	3.3	580073	4 US-08-545-528D-1	Sequence 1, Appl
C 19	34.6	3.2	618	4 US-09-248-796A-1121	Sequence 1121, Ap
C 20	34.6	3.2	1100	3 US-08-642-807A-32	Sequence 32, Appl
C 21	34.4	3.2	2007	4 US-09-614-221A-370	Sequence 370, App
C 22	34.2	3.2	717	3 US-08-998-416-602	Sequence 602, App
C 23	33.6	3.1	703	3 US-08-998-416-178	Sequence 178, App
24	33.6	3.1	723	3 US-08-998-416-952	Sequence 952, App
C 25	33.6	3.1	2139	4 US-09-248-796A-5542	Sequence 5542, Ap
C 26	33.6	3.1	2733	4 US-09-107-532A-2699	Sequence 2699, Ap
C 27	33.4	3.1	534	4 US-09-248-796A-1197	Sequence 1197, Ap

28	33.4	3.1	1068	4 US-09-328-352-2430	Sequence 2430, Ap
29	33.4	3.1	1167	3 US-09-308-003-6	Sequence 6, Appli
30	33.4	3.1	1539	4 US-08-956-171E-562	Sequence 562, App
31	33.4	3.1	1539	4 US-08-781-986A-562	Sequence 562, App
C 32	33.4	3.1	5163	3 US-08-700-651-1	Sequence 1, Appli
C 33	33.4	3.1	5163	3 US-08-928-361B-4	Sequence 4, Appli
C 34	33.4	3.1	5163	3 US-09-588-995A-4	Sequence 4, Appli
C 35	33.4	3.1	5318	3 US-08-700-651-2	Sequence 2, Appli
C 36	33.4	3.1	5318	3 US-08-928-361B-3	Sequence 3, Appli
C 37	33.4	3.1	5318	4 US-09-588-995A-3	Sequence 3, Appli
C 38	33.2	3.1	1233	3 US-08-858-207A-120	Sequence 120, App
39	33.2	3.1	1863	1 US-09-198-284-4	Sequence 4, Appli
40	33.2	3.1	1863	2 US-08-987-122-4	Sequence 4, Appli
C 41	33.2	3.1	3010	4 US-08-961-527-25	Sequence 25, Appl
C 42	33.2	3.1	3534	3 US-09-134-001C-2269	Sequence 2269, Ap
C 43	33.2	3.1	5191	4 US-09-809-665A-41	Sequence 41, Appl
C 44	33	3.1	3259	3 US-09-318-448-23	Sequence 23, Appl
C 45	33	3.1	8607	4 US-10-204-708-71	Sequence 71, Appl

ALIGNMENTS

RESULT 1  
US-08-232-463-14  
; Sequence 14, Application US/08232463  
; Patent No. 5670367  
; GENERAL INFORMATION:  
; APPLICANT: DORNER, F.  
; APPLICANT: SCHEIFLINGER, F.  
; APPLICANT: FALKNER, F. G.  
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Hardner  
; STREET: 1800 Diagonal Road, Suite 500  
; CITY: Alexandria  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22313-0299  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/232,463  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/935,313  
; FILING DATE:  
; APPLICATION NUMBER: EP 91 114 300.6  
; FILING DATE: 26-AUG-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BENT, Stephen A.  
; REGISTRATION NUMBER: 29,768  
; REFERENCE/DOCKET NUMBER: 30472/114 IMM  
; TELEPHONE: (703)836-9300  
; TELEFAX: (703)683-4109  
; TELEX: 899149  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7218 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; CLONE: pTZapt-F1s

Query Match 5.6%; Score 59.4; DB 1; Length 7218;







```
Matches 119; Conservative 0; Mismatches 136; Indels 0; Gaps 0;
Qy 308 TATAGCTTATCTTCAGTTCTTACAGCTGCTCTTATCATCTGTTACAGTTAGCTTTAT 367
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
632 TGTGTTGAGCTTGTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 573
Qy 368 AGCTATCTTCTCATCTTCTCATGTTTAAACATAAAGTTCACTCTTTTACCACCAATGCTGT 427
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
572 GAATAAGCTCTCACAGAAATGGTTGGCTTTTGAATAGATTGCTGATTAATAATATAT 513
Qy 428 TGTGTTGTTGACTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 487
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
512 TGTGTTGAAATGCTGTTGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 453
Qy 488 TCATGAGACTCACAGAGTACATAAATGCTGTTTCTTGATTACTGTAGCAGCAGCTGTTAT 547
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
452 TGAATTTGTTGAATGGTGTAAAGGTTGCTGCTCTCAATCTTGGCAACGGTAATGCTGAT 393
Qy 548 GTATGCTTTCATCTT 562
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
392 TGGGATGGGATCTT 378

RESULT 7
US-09-806-708B-23
; Sequence 23, Application US/09806708B
; Patent No. 6784342
; GENERAL INFORMATION:
; APPLICANT: The University of British Columbia
; TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
; FILE REFERENCE: 4810-58741
; CURRENT APPLICATION NUMBER: US/09/806,708B
; CURRENT FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/147,133
; PRIOR FILING DATE: 1999-08-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 23
; LENGTH: 1055
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (1)..(1055)
; OTHER INFORMATION: consensus sequence of A.t. and L.a. FAEI promoters
US-09-806-708B-23

Query Match 3.5%; Score 37.2; DB 4; Length 1055;
Best Local Similarity 18.6%; Pred. No. 0.37;
Matches 76; Conservative 124; Mismatches 205; Indels 3; Gaps 1;
Qy 204 GTGGGTGATGACAGTTTCTTCTTATCAACCGCGCTCTTCTTATGCGCGCTGTTATT 263
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
18 GWTNNTGRTTGGTGTWYCANNTGKCYARRWGMKTAYYMTATKMGTTGWAARTWR 77
Qy 264 GTAGCATCTCTCAGGTTTGTAACTACTGTATGATGATGATGATGATGATGATCTTCCA 323
Db ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
78 AAMKTRKWMCTAMNNNAWTCTARKWRTGTRWTKTNNAATGTRWWTGYM---TNNN 134
Qy 324 GTTTCACAGCTGCTCTTATCATCTTCAAGTATGCTTTTATAGCTATCTCTCATTC 383
Db ||||| ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
135 GCSITWARRYKTRRWYCYTAMWYGASWAGNASTRTTYTWKWMCKRKSARARATGR 194
Qy 384 TTATGTTTAAATAAAGTTCACCTCTTTTACCATCAATGCTGTGTGTTGTTGCTGTT 443
Db ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
195 ARYNRAWTARWRTGWTKAYAAWTMNNNNNAKAAKRAATTWGRKSNCTCTTAGG 254
Qy 444 GGTGCTGGGTTTGGGAATGCATACCGAACTGATAAGCCAGTTTCATGACCTACAG 503
Db ||||| ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
255 TTTKRAATCCWATCGAGWATKKWTKTSAAMGWTMNNNNNNNTTTTAAAYAAARWMMW 314
Qy 504 CAGTACATAAAGTGGTTTCTTGATTACTGTAGCAGCAGCTGTTATGATGCTTTCATCTG 563
Db :||| :||| ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: |||
```

```
Db 315 SATTTWAAAWTSRKTTWYYGRKTAAMNNNGTTCTTRMWAATWKMWMKTKGTTTNNNGG 374
Qy 564 CCATTAGTGAACCTTGCTTACCAGAAAGCTTAAGCAAAACCATGAGCTAT 611
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
375 RTVTGWTTKQVATTTTAKANNCTTAAWKWKTKTMMNTTAAKATTYAT 422

RESULT 8
US-08-916-421B-1/c
; Sequence 1, Application US/08916421B
; Patent No. 6503729
; GENERAL INFORMATION:
; APPLICANT: Bult et al.
; TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus
; Patent No. 6503729
; TITLE OF INVENTION: jannaschii
; FILE REFERENCE: PB275
; CURRENT APPLICATION NUMBER: US/08/916,421B
; CURRENT FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: US 60/024,428
; PRIOR FILING DATE: 1996-08-22
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1664976
; TYPE: DNA
; ORGANISM: Methanococcus jannaschii
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (28222)..(28222)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (28257)..(28258)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (84773)..(84773)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (84808)..(84808)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (84812)..(84812)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98120)..(98120)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98159)..(98159)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98239)..(98239)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98266)..(98266)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98343)..(98343)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (103998)..(103998)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (148948)..(148948)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (163385)..(163385)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (191989)..(191989)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (191995)..(191995)
; OTHER INFORMATION: n equals a, t, c, or g
```

```
/ NAME/KEY: misc feature
/ LOCATION: (231980)..(231980)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (234187)..(234187)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (234220)..(234220)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (234814)..(234814)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (309398)..(309398)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (309418)..(309418)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (312837)..(312837)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (312993)..(312993)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (319226)..(319226)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (559167)..(559167)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (559241)..(559241)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (600992)..(600992)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (622708)..(622708)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (657081)..(657081)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (657203)..(657203)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (674435)..(674435)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (682442)..(682442)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (713652)..(713652)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (741684)..(741684)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (779455)..(779455)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (779676)..(779676)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (855539)..(855539)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (871619)..(871619)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (1084830)..(1084830)
/ OTHER INFORMATION: n equals a, t, c, or g
```

```
/ LOCATION: (1096846)..(1096846)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (1119881)..(1119881)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (1130881)..(1130881)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (1310988)..(1310988)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (1313224)..(1313224)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (1349473)..(1349473)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (1470091)..(1470091)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (1569020)..(1569020)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (1602912)..(1602912)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (1603734)..(1603734)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (1637998)..(1637998)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (1664854)..(1664854)
/ OTHER INFORMATION: n equals a, t, c, or g
/ US-08-916-421B-1

Query Match 3.5%; Score 37.2; DB 4; Length 1664976;
Best Local Similarity 43.8%; Pred. No. 19;
Matches 162; Conservative 0; Mismatches 208; Indels 0; Gaps 0;

Qy 213 AGTACAAGTTCTTTCTTATCAACCGGTCTCTTATCGCGCTGTTATTCTAGGCATT 272
Db 1077557 ACTACTGGATTGGTATTTAAAGAGAGTATTTATAAGCATCTGCAGTTATTGGGTTAATT 1077498

Qy 273 CTCTCAGGGTTTGATPAACTCTTGTATGCATATGGTATAGCTTATCTTCCAGTTTCTACA 332
Db 1077497 CTATTGTTGTTATTTTATTCGTGTTTTTACAGCATATAACTCAITTTGGAAGCAATTTTA 1077438

Qy 333 GCTGCTTTATCATGCTTCACAGTTAGCTTTTATAGCTATCTTCTCATCTTCATGGTT 392
Db 1077437 TACGTTGGAATCTTAGTTTGAATATTTTATTATTAGTAAATTTTATAGATAAAACCTTA 1077378

Qy 393 AACATAAAGTTTCACCTCCTTTTACCACATCAATGCTGTTGTTGTTGTTGCTGCTGG 452
Db 1077377 AAATTAATATCATACAGTCAAGATTGATGATTAATGTTGGCGCAGTTGATTGAT 1077318

Qy 453 GTTTTGGGAATGCATACCAGAACTGATAAGCCAGTTTCATGAGACTCACAAAGCAGTACATA 512
Db 1077317 GCTTCTGCAACAACATATTGGAATAGGAGTTTATGTTTATTTGGAACAACATCCA 1077258

Qy 513 ACTGGTTTCTTGATTACTGTAGCAGCAGCTGTTATGTTGTTTCTCATCTTGCATATGTTG 572
Db 1077257 AGATTTTGTGGAAGAACTTTTGGAGTTTATGCTTTTATTCCTTTTAAAGCTCTTAATAATT 1077198

Qy 573 GAACCTTGCTT 582
Db 1077197 TTATTAGCTT 1077188
```

```
RESULT 9
US-09-692-570-1/c
; Sequence 1, Application US/09692570
; Patent No. 6797466
; GENERAL INFORMATION:
; APPLICANT: Bult et al.
; TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus
; TITLE OF INVENTION: jannaschii
; FILE REFERENCE: PB275CI
; CURRENT APPLICATION NUMBER: US/09/692,570
; CURRENT FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: US 60/024,428
; PRIOR FILING DATE: 1996-08-22
; PRIOR APPLICATION NUMBER: US 08/916,421
; PRIOR FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1664976
; TYPE: DNA
; ORGANISM: Methanococcus jannaschii
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (28222)..(28222)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (28257)..(28258)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (84773)..(84773)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (84808)..(84808)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (84812)..(84812)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (98120)..(98120)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (98159)..(98159)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (98239)..(98239)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (98266)..(98266)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (98343)..(98343)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (103998)..(103998)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (148948)..(148948)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (163385)..(163385)
; OTHER INFORMATION: n equals a, t, c, or g
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (191989)..(191989)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (191995)..(191995)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (231980)..(231980)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (234187)..(234187)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (234220)..(234220)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (234814)..(234814)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (309398)..(309398)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (309418)..(309418)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (312837)..(312837)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (312993)..(312993)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (319226)..(319226)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (559167)..(559167)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (559241)..(559241)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (600992)..(600992)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (622708)..(622708)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (657081)..(657081)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (657203)..(657203)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (674435)..(674435)
; OTHER INFORMATION: n equals a, t, c, or g
```



;  
; CITY: Wilmington  
; STATE: Delaware  
; COUNTRY: U.S.A.  
; ZIP: 19899  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 inch diskette  
; COMPUTER: IBM/PC or Compatible  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: Wordperfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/973,257  
; FILING DATE: 19921109  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA: No. 5378820e  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4182 base pairs  
; TYPE: NUCLEIC ACID  
; STRANDEDNESS: double stranded  
; TOPOLOGY: linear  
; MOLECULE TYPE: Genomic DNA  
; HYPOTHETICAL: No  
; ANTI-SENSE: No  
; ORIGINAL SOURCE:  
; ORGANISM: Mycoplasma Gallisepticum  
; STRAIN: S6  
; US-07-973-257-1

Query Match 3.5%; Score 37; DB 1; Length 4182;  
Best Local Similarity 53.9%; Pred. No. 0.9;  
Matches 76; Conservative 0; Mismatches 65; Indels 0; Gaps 0;  
  
QY 319 TTCAGTTCTACAGCTGCTTATCAATGCTTACAGTACCTTTTATAGTATCTTCT 378  
DB 4163 TTCCAAATACAGCTAATCCACTATCAATGATATACTTTTAAAGTTTTTTAGAAATATCA 4104  
  
QY 379 CATTTCTTCAATGTTAAACATAAAGTTCACTCTCTTTTACCATCAATGCTGTGTGTTGTA 438  
DB 4103 TAATCTTAATATATACCTTAATATTTCTTCTAGTTGGTCTGGAGCTTTTGTT 4044  
  
QY 439 CTGTTGGTCTGCGGTTTTGG 459  
DB 4043 TAGCTGGTCTGTTGGTTTTG 4023

RESULT 12  
US-09-202-712-18  
; Sequence 18, Application US/09202712  
; Patent No. 6265637  
; GENERAL INFORMATION:  
; APPLICANT: Coupland, George M  
; APPLICANT: Schaffer, Robert J  
; TITLE OF INVENTION: Genetic control of flowering  
; FILE REFERENCE: 620-55  
; CURRENT APPLICATION NUMBER: US/09/202,712  
; CURRENT FILING DATE: 1999-01-11  
; EARLIER APPLICATION NUMBER: PCT/GB97/01676  
; EARLIER FILING DATE: 1997-06-23  
; EARLIER APPLICATION NUMBER: GB 9613132.1  
; EARLIER FILING DATE: 1996-06-21  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 18  
; LENGTH: 1738  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-09-202-712-18

Query Match 3.4%; Score 36.2; DB 3; Length 1738;  
Best Local Similarity 53.1%; Pred. No. 0.98;  
Matches 77; Conservative 0; Mismatches 68; Indels 0; Gaps 0;  
  
QY 212 TAGTACAAGTTTCTTTCTTATCAACCGCGTCTTCTTATCGCGCTGTTATTGTAGGCAT 271

DB 1018 TAAACAGCTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1077  
QY 272 TCTCTCAGGGTTTGAATAACTACTTGTATGATATGATAGCTTATCTTCCAGTTTCTTAC 331  
DB 1078 TCTCTCTGTTTTTTTAAATTTATTTTTTTAGAGATTTTTTTTTTGTGTTCCGATTGATT 1137  
QY 332 AGCTGCTCTTATCATCTTCTTCCACAG 356  
DB 1138 ATTTCCGGGAACGATGACTTCTCTCCG 1162

RESULT 13  
US-09-621-976-8976/c  
; Sequence 8976, Application US/09621976  
; Patent No. 6639063  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Jobert, S.  
; APPLICANT: Giordano J.Y.  
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
; FILE REFERENCE: GENSET.054PR2  
; CURRENT APPLICATION NUMBER: US/09/621.976  
; CURRENT FILING DATE: 2000-07-21  
; NUMBER OF SEQ ID NOS: 19335  
; SOFTWARE: Patent.pm  
; SEQ ID NO 8976  
; LENGTH: 399  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-621-976-8976

Query Match 3.4%; Score 35.8; DB 4; Length 399;  
Best Local Similarity 12.5%; Pred. No. 0.59;  
Matches 35; Conservative 134; Mismatches 106; Indels 6; Gaps 1;  
  
QY 669 ATCCGTATGTTCACTCGCTGGTGATTTCAGCAGCGCTTACAAAGAGCAAGAGAGTTC 728  
DB 314 AKMKRKKTKMYMKGGKGGSTYAMRSRGSTGRWSYRRAMWRGSKSWGGSYRM 255  
QY 729 AAGCTTCGAGAGGCAATGTTTCTATGTGTGCTGTGTTTTCAGCC-----ATCATATGG 782  
DB 254 AGYRSSRWSYSAMWRKKMTCKWGRSSWGRSTGYAWMYKSWCTSRKMYKKRRK 195  
QY 783 CAAGGCTTCTTCTGGGAGCCATTTGATTAATCTTCTCCACATCGTCTCGTCTCGGT 842  
DB 194 KWRKCTSTRTCYRGSTYKCAAYTKGRKWRWTTTTYKYSMSMKTKRMKTAYYWT 135  
QY 843 ATTATGATATCAGTGTCTTTTCCCAATTACAGAGGTTTTAGCTGTATATTTACCATGAA 902  
DB 134 KRWKTRTKWTCTMCWKCTTYWAGTMVYRYYAKRWSKRCCTWSTTTCYCKMYM 75  
QY 903 AAGTTTCAAGTGAAGAGGACTTTCTTCTGCTCTCTCCCT 943  
DB 74 AKCWSYMSMSMMKMGKSMWMTYTTTTYMMKWSKMTY 34

RESULT 14  
US-09-621-976-2813  
; Sequence 2813, Application US/09621976  
; Patent No. 6639063  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Jobert, S.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
; FILE REFERENCE: GENSET.054PR2  
; CURRENT APPLICATION NUMBER: US/09/621.976  
; CURRENT FILING DATE: 2000-07-21  
; NUMBER OF SEQ ID NOS: 19335  
; SOFTWARE: Patent.pm  
; SEQ ID NO 2813  
; LENGTH: 832

```
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 235..399
US-09-621-976-2813

Query Match
Best Local Similarity 3.4%; Score 35.8; DB 4; Length 832;
Matches 25; Conservative 103; Mismatches 85; Indels 0; Gaps 0;

QY 816 TTCTCCACATCGTCTCTCGTCTCGGTATATGATATAGTGTCTTTTGGCAATTACAGAG 875
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
7 KYTTWAKCWTWKWSWSYMYWKYMKTYWRERKKKAWKWKYKWTWTWYRYAMWG 66

QY 876 GTTTTAGCTGTTATATTCTACCATGAAAGTTTCAAGCTGAGAAAGGACTTTCTCTTGCT 935
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
67 TYKKKAMCRTKTKKKKKKGYMMWYWGWRRSYAMWTRTWGYAYRSMYMWRYRCWKK 126

QY 936 CTCCTCCCTTTGGGGCTTTGTCTCTTACTTTTATGGTGAGATAAGTCTGGCGAGGATAAA 995
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
127 KAYRKTTTCYSSKGTWWRKWKAWTTWWKKTYWAAATRYMMWMCWTKWRASWYCWY 186

QY 996 AGGAGAAATTCAGCAGGAGAGAGTCTAGGAGACA 1028
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
187 WGRKRWSTWRKRSYASARSARCCYSCSWG 219

RESULT 15
US-09-248-796A-751/c
; Sequence 751, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 751
; LENGTH: 534
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-751

Query Match
Best Local Similarity 3.3%; Score 35.4; DB 4; Length 534;
Matches 84; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

QY 219 AGTTCTCTTTCTATCAACCGCTCTTCTTATCGCGCGTGTATTTGAGGCATTTCTCTCA 278
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
518 AATTGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 459

QY 279 GGGTTTGATAACTACTTGTATGATATGGTATAGCTTATCTCCAGTTTCTACAGTGTCT 338
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
458 GCCAATCTTTTCATACCTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 399

QY 339 CTTATCATTTGCTTCACAGTTAGCTTTTATAGCTATCTTCTCATTC 383
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
398 ATTCGTTGATTTCTAGCAGATTTCTTTTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 354

Search completed: November 1, 2004, 22:41:08
Job time : 104.435 secs
```

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 1, 2004, 05:45:29 ; Search time 4253.75 Seconds  
(without alignments)  
10228.408 Million cell updates/sec

Title: US-09-913-767-5  
Perfect score: 1194  
Sequence: 1 tcatgagatataataacat.....tgtgatcaagcatatttcc 1194

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST: \*  
1: gb\_est1: \*  
2: gb\_est2: \*  
3: gb\_hc: \*  
4: gb\_est3: \*  
5: gb\_est4: \*  
6: gb\_est5: \*  
7: gb\_est6: \*  
8: gb\_gsa1: \*  
9: gb\_gsa2: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1151.6	96.4	1340	CNS0A29J	BX827536 Arabidops
2	1091.2	91.4	1359	CNS0A2R3	BX827477 Arabidops
3	941.6	78.9	1431	CNS0A2KR	BX829081 Arabidops
4	909.2	76.1	1375	CNS0A2S8	BX827187 Arabidops
5	903	75.6	1312	CNS0A2U8	BX827470 Arabidops
C 6	442.6	37.1	709	BH498028	B77316 T3213TF TAM
7	442.6	37.1	709	BH498028	BH498028 BOHQK68TF
8	437.4	36.6	691	BZ087060	BZ087060 lkh10d12
9	396	32.3	558	AV828990	AV828990 AV828990
C 10	334.8	28.0	446	AI996743	AI996743 701688184
11	317	26.5	1519	CNS0A3B5	BX826795 Arabidops
12	312.8	26.2	1323	CNS0A36P	BX826811 Arabidops
C 13	299.6	25.1	498	AU226302	AU226302 AU226302
14	239.8	20.1	922	BZ967767	BZ967767 PUDGUS6TD
C 15	235.8	19.9	428	AU227210	AU227210 AU227210
C 16	235.8	19.7	424	BP575101	BP575101 BP575101
17	234.8	19.7	739	CO105307	CO105307 GR_Eb003
18	231.2	19.4	568	BM176948	BM176948 saJ74h01
C 19	229.8	19.2	782	CO105310	CO105310 GR_Eb003
C 20	228.6	19.1	770	CG344371	CG344371 OG4AG15TC
C 21	228.4	19.1	822	BZ491815	BZ491815 BONFU42TF
C 22	224.6	18.8	837	CB292743	CB292743 UCRC501_0
C 23	224.4	18.8	351	AV441569	AV441569 AV441569
C 24	224.2	18.8	444	AV439952	AV439952 AV439952

25	220	18.4	806	4	BI310188	EST531193
26	219.6	18.4	666	5	BQ157452	NF105A121
27	219.2	18.4	1374	3	CNS0ABEV	BX814955 Arabidops
28	213	17.8	401	7	T42237	5500 Lambda
29	212.6	17.8	627	5	BQ155273	NF078F121
30	212.6	17.8	729	8	BH470210	BOHAN77TR
31	211.4	17.7	905	5	BX927581	BX927581
32	208.2	17.4	768	1	AJ795195	AJ795195
33	208	17.4	780	6	CD485871	CFUS4.1C1
34	200.4	16.8	633	1	AI777352	EST263760
C 35	199.8	16.7	426	5	BP666513	BP666513
C 36	198.4	16.6	562	7	CK759270	pam01-13m
37	195	16.3	668	5	BQ704880	Bn01_01n0
38	190.6	16.0	630	4	BG525645	52-8 Stev
39	190.6	16.0	762	5	BQ121182	EST606758
40	188.2	15.8	882	7	CN125340	RHO1_10
41	186.6	15.6	652	6	CA255028	SCEPFL417
42	184.6	15.5	450	8	BH509846	BOHHR68TR
43	183.8	15.4	733	6	CB292744	UCRC501_0
44	183.6	15.4	830	8	BZ500536	BONFV75TF
C 45	182.4	15.3	735	4	BM408950	EST583277

#### ALIGNMENTS

RESULT 1  
CNS0A29J  
LOCUS  
DEFINITION  
CNS0A29J 1340 bp mRNA linear HTC 06-FEB-2004  
Arabidopsis thaliana Full-length cDNA Complete sequence from clone  
GSLTSL712A05 of Adult vegetative tissue of strain col-0 of  
Arabidopsis thaliana (thale cress).  
ACCESSION  
BX827536.1 GI:42459416  
VERSION  
HTC; GSLT\_cDNA.  
KEYWORDS  
Arabidopsis thaliana (thale cress)  
SOURCE  
Arabidopsis thaliana  
ORGANISM  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
REFERENCE  
1 (bases 1 to 1340)  
Castelli, V., Aury, J.M., Jaillon, O., Wincker, P., Clepet, C.,  
Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V.,  
Temple, G., Caboche, M., Weissenbach, J., and Salanoubat, M.  
Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences:  
A Combined Approach to Evaluate and Improve Arabidopsis Genome  
Annotation  
Unpublished  
2 (bases 1 to 1340)  
Genoscope.  
Direct Submission  
Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :  
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr  
- Web : www.genoscope.cns.fr)  
The sequences are based on single pass reads.  
Life Technologies (a division of Invitrogen)  
full-length libraries construction : Temple G  
Genoscope members carried out sequencing and annotation : Castelli  
V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,  
Schachter V., Weissenbach J., Salanoubat M.  
URGV INRA : Clepet C., Caboche M.  
Annotation is based on the June 2003 version of the Arabidopsis  
genome released by MIPS (Munich Information center for Protein  
Sequences). 5 prime and 3 prime are assembled with Phrap:  
http://www.genoscope.cns.fr/externe/sequences/Banque\_Projet\_EF/Full  
length  
http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.  
Location/Qualifiers  
1. 1340  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/strain="Col-0"  
/db\_xref="taxon:3702"

JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
Unpublished  
2 (bases 1 to 1340)  
Genoscope.  
Direct Submission  
Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :  
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr  
- Web : www.genoscope.cns.fr)  
The sequences are based on single pass reads.  
Life Technologies (a division of Invitrogen)  
full-length libraries construction : Temple G  
Genoscope members carried out sequencing and annotation : Castelli  
V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,  
Schachter V., Weissenbach J., Salanoubat M.  
URGV INRA : Clepet C., Caboche M.  
Annotation is based on the June 2003 version of the Arabidopsis  
genome released by MIPS (Munich Information center for Protein  
Sequences). 5 prime and 3 prime are assembled with Phrap:  
http://www.genoscope.cns.fr/externe/sequences/Banque\_Projet\_EF/Full  
length  
http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.  
Location/Qualifiers  
1. 1340  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/strain="Col-0"  
/db\_xref="taxon:3702"

FEATURES  
source

/clone="GSLTSL71ZA05"	
/tissue type="Adult vegetative tissue"	
/plasmid="pCMVSPORT_6"	
1. .1340	
/gene="At4g18197"	
gene	
ORIGIN	
Query Match	96.4%; Score 1151.6; DB 3; Length 1340;
Besc Local Similarity	99.7%; Pred. No. 0;
Matches 1154; Conservative	0; Mismatches 4; Indels 0; Gaps 0;
Qy	37 GGTGACCAAGAACTTAGAAGCAAACTTATAGATCATGAGGTGGTAACCTGAATCATCATCA 96
Db	65 GGTGACCAAGAACTTAGAAGCAAACTTATAGATCATGAGGTGGTAACCTGAATCATCATCA 124
Qy	97 TCAGCTGTGCCTCAAAACCGAGAACTATATAAAAGGTGGCTTCGTGTCTCCATATACGTAATC 156
Db	125 TCAGCTGTGCCTCAAAACCGAGAACTATATAAAAGGTGGCTTCGTGTCTCCATATACGTAATC 184
Qy	157 TTTGTCCCTTTTGGCCAGCCACTAGCTACAATTTCTGGGTAGATTGTACTATGAATGGA 216
Db	185 TTTGTCCCTTTTGGCCAGCCACTAGCTACAATTTCTGGGTAGATTGTACTATGAATGGA 244
Qy	217 GGAATAGCACATATGTGGTAACACTTCTTCAACTCATTTGGCTTCCCTGTACTGGTTCGT 276
Db	245 GGAATAGCACATATGTGGTAACACTTCTTCAACTCATTTGGCTTCCCTGTACTGGTTCGT 304
Qy	277 TTCGCTCTTTTTCGAAATCAGGCAACCCAAATCAACAGATACAAAATTCAGTCAGTCC 336
Db	305 TTCGCTCTTTTTCGAAATCAGGCAACCCAAATCAACAGATACAAAATTCAGTCAGTCC 364
Qy	337 CCTTCCTTCACACCCCTTGATCGGTTTACTTGTGCACTGCAGCTAGTTCGCTTAT 396
Db	365 CCTTCCTTCACACCCCTTGATCGGTTTACTTGTGCACTGCAGCTAGTTCGCTTAT 424
Qy	397 GCTTATTTGTCTGCAGTAGGGTTGCTCTACTTACCACTCTACTTTTCCCTCATCTTG 456
Db	425 GCTTATTTGTCTGCAGTAGGGTTGCTCTACTTACCACTCTACTTTTCCCTCATCTTG 484
Qy	457 GCCTCACAGTTGGCCTTCACCTGCTTTTTCATATTTTCTTAACTCGCAAAAGTTCACT 516
Db	485 GCCTCACAGTTGGCCTTCACCTGCTTTTTCATATTTTCTTAACTCGCAAAAGTTCACT 544
Qy	517 CCTTTGATGATCAGTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 576
Db	545 CCTTTGATGATCAGTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 604
Qy	577 ACTGATTCAGAAAACTCAACTAATGTATCTAGAGTACAGTATGTGATCGGTTTCATATGT 636
Db	605 ACTGATTCAGAAAACTCAACTAATGTATCTAGAGTACAGTATGTGATCGGTTTCATATGT 664
Qy	637 ACCATCGGTGTTCCGCTGGGATGGAGCTGTATCTATCTGTATACAAATGCTCTTCAGG 696
Db	665 ACCATCGGTGTTCCGCTGGGATGGAGCTGTATCTATCTGTATACAAATGCTCTTCAGG 724
Qy	697 AAAGTTTTCAGAGCATACATCTCAGACAGTACGGACTTGGCCATTTACAGTCTCTA 756
Db	725 AAAGTTTTCAGAGCATACATCTCAGACAGTACGGACTTGGCCATTTACAGTCTCTA 784
Qy	757 GTTGGAGTTGTGTAGTTCTCATAGGACTTTTTCGAAGTGGAGAGTGGAACTTTTGCCA 816
Db	785 GTTGGAGTTGTGTAGTTCTCATAGGACTTTTTCGAAGTGGAGAGTGGAACTTTTGCCA 844
Qy	817 AGTGAGATGAGAACTACAACTCGGGAAGTGTATATGTTTGTGACTTTACGCTTCGGCA 876
Db	845 AGTGAGATGAGAACTACAACTCGGGAAGTGTATATGTTTGTGACTTTACGCTTCGGCA 904
Qy	877 GCTATTTCTCGCAGTCTACACTCTGCTTGTGGGATGATCTTCGAGTCACTCTCT 936
Db	905 GCTATTTCTCGCAGTCTACACTCTGCTTGTGGGATGATCTTCGAGTCACTCTCT 964
Qy	937 GTGTTCTCCAATTCATAACAGCTGTGGGATTCCTATAGTTCCAGTTGCGGCAGTGATA 996
Db	965 GTGTTCTCCAATTCATAACAGCTGTGGGATTCCTATAGTTCCAGTTGCGGCAGTGATA 1024
Qy	997 GTTTTTCATGATAGATGGAGCATCCAAAATCTTCTCCATTTATTTTAGCTATCTGGGC 1056
Db	1025 GTTTTTCATGATAGATGGAGCATCCAAAATCTTCTCCATTTATTTTAGCTATCTGGGC 1084
Qy	1057 TTTCTTTTCATTTCTGCTATCAGCACTACTCTCGAGAAAGAGTTGTAATCTAGGCCACACA 1116
Db	1085 TTTCTTTTCATTTCTGCTATCAGCACTACTCTCGAGAAAGAGTTGTAATCTAGGCCACACA 1144
Qy	1117 AGTGCTGTAGGAGATCTTCTATCTACCTGTGGAGAGGTCCACACAAACATACAAAGTGTG 1176
Db	1145 AGTGCTGTAGGAGATCTTCTATCTACCTGTGGAGAGGTCCACACAAACATACAAAGTGTG 1204
Qy	1177 TGATCAAGCATATTTTCC 1194
Db	1205 TGATCAAGCATATTTTCC 1222
RESULT 2	
CNS0A2R3	
LOCUS	1359 bp mRNA linear HTC 06-FEB-2004
DEFINITION	Arabidopsis thaliana Full-length cDNA Complete sequence from clone GSLTSL552B07 of Adult vegetative tissue of strain col-0 of Arabidopsis thaliana (thale cress).
ACCESSION	BX827477.1 GI:42459927
VERSION	HTC; GSLT_cDNA.
KEYWORDS	Arabidopsis thaliana (thale cress)
SOURCE	Arabidopsis thaliana
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis. 1 (bases 1 to 1359)
REFERENCE	Castelli,V., Aury,J.M., Jaillon,O., Wincker,P., Clepet,C., Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V., Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M. Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences: A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation
AUTHORS	Annotation Unpublished 2 (bases 1 to 1359)
TITLE	Genoscope.
JOURNAL	Direct Submission
REFERENCE	Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
AUTHORS	The sequences are based on single pass reads. Life Technologies (a division of Invitrogen) members carried out full-length libraries construction : Temple G. Genoscope members carried out sequencing and annotation : Castelli V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M.
JOURNAL	URGV INRA : Clepet C., Caboche M.
COMMENT	Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich Information center for Protein Sequences). 5 prime and 3 prime are assembled with Phrap. http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EP/Full length http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis. Location/Qualifiers 1. .1359 /organism="Arabidopsis thaliana" /mol_type="mRNA" /strain="Col-0" /db_xref="taxon:3702" /clone="GSLTSL552B07" /tissue type="Adult vegetative tissue" /plasmid="pCMVSPORT_6" 1. .1359 /gene="At4g18197"
FEATURES	
source	
gene	
ORIGIN	

Query Match		91.4%;	Score 1091.2;	DB 3;	Length 1359;
Best Local Similarity		97.4%;	Pred. No. 2.1e-301;		
Matches 1130;		Conservative	0;	Mismatches 28;	Indels 2; Gaps 2;
QY	37	GGTGACCAAGCTTGAAGCAAACTTATAGATCATGAGGTGGTAACTGAATCATCATCA	96		
Db	39	GGTGACCAAGCTTGAAGCAAACTTATAGATCATGAGGTGGTAACTGAATCATCATCA	98		
QY	97	TCAGCTGTGCTCAAAACCGAGAACTATAAAGGTGGCTTCGTGTCTCCATATACGTAATC	156		
Db	99	TCAGCTGTGCTCAAAACCGAGAACTATAAAGGTGGCTTCGTGTCTCCATATACGTAATC	158		
QY	157	TTTGCTCTCTTTTGCAGGCACTAGCTACAAATTCGGGTAGATGTGTAATGAATGA	216		
Db	159	TTTGCTCTCTTTTGCAGGCACTAGCTACAAATTCGGGTAGATGTGTAATGAATGA	218		
QY	217	GGAAATAGACATATGTGTACACCTTCTCAACTCATTTGGCTCCCTGACTGTTCTCG	276		
Db	219	GGAAATAGACATATGTGTACACCTTCTCAACTCATTTGGCTCCCTGACTGTTCTCG	278		
QY	277	TTCCGCTTCTTTTCTCGAATCAGGCAACCCAAATCAACAGATACAAATTTCACTCAGTCC	336		
Db	279	TTCCGCTTCTTTTCTCGAATCAGGCAACCCAAATCAACAGATACAAATTTCACTCAGTCC	338		
QY	337	CTTCTCTTCAACACCTTGTGATCGGTTTACTTGTGCACTGGACTGTAGTTCGCTTAT	396		
Db	339	CTTCTCTTCAACACCTTGTGATCGGTTTACTTGTGCACTGGACTGTAGTTCGCTTAT	398		
QY	397	GCTTATTTGTCTGCACTAGGTTGCTCTACTTACCAGTCTCTACTTTCTCCCTCATCTTG	456		
Db	399	GCTTATTTGTCTGCACTAGGTTGCTCTACTTACCAGTCTCTACTTTCTCCCTCATCTTG	458		
QY	457	GCTCACAGTTGGCTTCTCACTGCTTTTCTCATATTTCTTAACTCGCAAAAGTTCACT	516		
Db	459	GCTCACAGTTGGCTTCTCACTGCTTTTCTCATATTTCTTAACTCGCAAAAGTTCACT	518		
QY	517	CTTTGATAGTCACTTTCTCTCTCACTGATCTCTGCTCTTCTTGTGGTCAAC	576		
Db	519	CTTTGATAGTCACTTTCTCTCTCACTGATCTCTGCTCTTCTTGTGGTCAAC	578		
QY	577	ACTGATTCAGAACTCACTAATGTATGTAGTACAGTATGTATCGGTTCTATGT	636		
Db	579	ACTGATTCAGAACTCACTAATGTATGTAGTACAGTATGTATCGGTTCTATGT	638		
QY	637	ACCATCGTGTCTCCGCTGGGATTTGACCTTACTATCTCTGATACAAATGTCTTCAGG	696		
Db	639	ACCATCGTGTCTCCGCTGGGATTTGACCTTACTATCTCTGATACAAATGTCTTCAGG	698		
QY	697	AAAGTTTTCAAGACATACCTCTCAGCAGTCAAGGCTTGGCCATTTACCAGTCTCTA	756		
Db	699	AAAGTTTTCAAGACATACCTCTCAGCAGTCAAGGCTTGGCCATTTACCAGTCTCTA	758		
QY	757	GTTCGAGTTGTAGTTCTCATAGACCTTTTTCAGTGGAGTGGGAACTTTGCCA	816		
Db	759	GTTCGAGTTGTAGTTCTCATAGACCTTTTTCAGTGGAGTGGGAACTTTGCCA	818		
QY	817	AGTGAGATGAGAACTACAAACTCGGAAAGTGTCTATGTTTGTAGCTTTAGCCTCGGCA	876		
Db	819	CGTGAGATGAGAACTACAAACTCGGAAAGTGTCTATGTTTGTAGCTTTAGCCTCGGCA	878		
QY	877	GCTATTTCTCGCAAGTCTACACTCTCTGCTTGTGGGAT-TGATCTTCG-AGTCATCT	934		
Db	879	ACTATTTTCATGGCAAGTCTACACTCTTGTCTTGTGGGATATGATCTTCGCACTCTCT	938		
QY	935	CTGTGTTCTCCAAATTCATTAACAGCTGTGGGATTTGCTATAGTTCCAGTTGGCGAGTGA	994		
Db	939	CTGTGTTTCACCAATTCATTCGCCGTGTGGGAGTGCCTATCTGTTCCGCTGGCGCGTGA	998		
QY	995	TAGTTTTTCATGATGAATGGAGCATCCAAATCTTCTCCATTATTTTAGCTATCTCGG	1054		
Db	999	TCGTTTTTCATGATGAATGGAGCATCAAAATCTTCTCCATTATTTTAGCTATCTCGG	1058		
QY	1055	GCTTCTCTTTCATTCGTCTATCAGCACTACCTCGACGAAAGAGTTGAATACTAGCCACA	1114		
Db	1059	GCTTCTCTCTCATTCGTCTATCAGAACTACCTCGACGAAAGAGTTGAATACTAGCCACA	1118		
QY	1115	CAAGTGCTGTAGGAGATCTTTCATCTACCTGTGTAGGAAAGGTCACACAACATACAAAGTG	1174		
Db	1119	CAAGTGCTGTAGGAGATCTTTCATCTACCTGTGTAGGAAAGGTCACACAACATACAAAGTG	1178		
QY	1175	TGTGATCAAAAGCATATTTTCC	1194		
Db	1179	TGTGATCAAAAGCATATTTTCC	1198		
RESULT 3					
CNSO42KR		1431 bp	mRNA	linear HTC 06-FEB-2004	
LOCUS		Arabidopsis thaliana Full-length cDNA Complete sequence from clone			
DEFINITION		GSLTSL1692B03 of Silique of strain col-0 of Arabidopsis thaliana (thale cress).			
ACCESSION		BX829081			
VERSION		BX829081.1 GI:42460228			
KEYWORDS		HTC; GSLT cDNA.			
SOURCE		Arabidopsis thaliana (thale cress)			
ORGANISM		Arabidopsis thaliana			
REFERENCE		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.			
AUTHORS		Castelli, V., Aury, J.M., Jaillon, O., Wincker, P., Clepet, C., Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V., Temple, G., Caboche, M., Weissenbach, J. and Salanoubat, M.			
TITLE		Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences: A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation			
JOURNAL		Unpublished			
REFERENCE		2 (bases 1 to 1431)			
AUTHORS		Genoscope.			
JOURNAL		Direct Submission			
TITLE		Submitted (18-NOV-2003) Genoscope - Centre National de Sequenage :			
JOURNAL		BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr			
COMMENT		- Web : www.genoscope.cns.fr)			
		The sequences are based on single pass reads.			
		Life Technologies (a division of Invitrogen) members carried out			
		full-length libraries construction : Temple G.			
		Genoscope members carried out sequencing and annotation : Castelli			
		V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,			
		Schachter V., Weissenbach J., Salanoubat M.			
		URGV INRA : Clepet C., Caboche M.			
		Annotation is based on the June 2003 version of the Arabidopsis			
		genome released by MIPS (Munich information center for Protein			
		Sequences) . 5 prime and 3 prime are assembled with Phrap.			
		http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full			
		length			
FEATURES		http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.			
source		Location/Qualifiers			
		1. .1431			
		/organism="Arabidopsis thaliana"			
		/mol_type="mRNA"			
		/strain="Col-0"			
		/db_xref="taxon:3702"			
		/clone="GSLTSL1692B03"			
		/issue_type="Silique"			
		/plasmid="pCMVSPORT_6"			
		1. .1431			
		/gene="At4gl8205"			
gene					
ORIGIN					
Query Match		78.9%;	Score 941.6;	DB 3;	Length 1431;
Best Local Similarity		90.4%;	Pred. No. 1.8e-258;		
Matches 1020;		Conservative	0;	Mismatches 99;	Indels 9; Gaps 1;
QY	3	ATGAGATATAATAACATGAGTGTGTTATTTTTCAGTGACCAAGAACTTAGAAGCAAACT	62		
Db	124	ATGAATATAATAACCAAGAGTGTGTTATTTTTCAGTGACCAAGAACTTAGAAGCAAACT	183		

```
QY 63 TATAGATCATGGGTGTTAACTGAATCATCATCATCAGCTGTGCTCAAAACCGAGAACTA 122
Db 184 TCTAGATCATGAG-----GAAACTGAATCATTTTCAGTACCTCAAGCAAGAACTG 234
QY 123 TAAAGGTGGCTTCGTGTCTCCATATACGTAATCTTTGTCTCTTTTGCCAGCAGCTAGC 182
Db 235 TAAGAGGTGGCTCGGTGTCTCCATATACGCATCTTTGTCACTCTTGCCAAACCACTGC 294
QY 183 TACAATCTGGGTAGATTGTACTATGAAATGGAGGAAATAGCACAATATGTTGGTAACT 242
Db 295 TACAGTTCTGGGTAGACTGTACTATGAAATGGAGGAAAGCAATATGTTGGTAACT 354
QY 243 TCTTCACTCATGCTCCGTCTACTGTTCTGTTCTGCTTCTTCTGCAATCAGGCA 302
Db 355 TCTTCACTCATGCTCCGTCTACTGATTCTCTTCCGCTTCTTTCTCGAATCAGGCA 414
QY 303 ACCCAATCAACAGATACAAATTTTCAGTCAGTCCCTTCTTCCACCCCTTGATCGGT 362
Db 415 ACCCAATCAACAGATACAAATTTTCAGTCAGTCCCTTCTTCCACCCCTTGATCGGT 474
QY 363 TTACTTGTGCACTGGACTGTAGTGTCCGTATATGCTTATTTGTCTGCACTGAGGTTGCT 422
Db 475 TTACTTGTGCACTGGACTGTAGTGTCTGCTTATGCTTATTTGTCTGCACTGAGGTTGCT 534
QY 423 CTACTTACAGTCTCTACTTTCTCCCTCATCTTGGCCCTCAGCTTGGCCCTTCACTGCTT 482
Db 535 TTACTTACAGTCTCTACTTTCTCCCTCATCTTGGCCCTCAGCTTGGCCCTTCACTGCTT 594
QY 483 TTTCTCATATTTCTTAACTCGAAAGTTCACATGCTTCTTGTAGTCACTTCTTGTCTTCT 542
Db 595 TTTCTCATATTTCTTAACTCGAAAGTTCACATGCTTCTTGTAGTCACTTCTTGTCTTCT 654
QY 543 CCTCAGTATCTCTGCTCTTCTTGTGGTCAACATGATTCAGAAAACTCAACTAATGT 602
Db 655 CCTTACGGTTTCTCTGCGCTCTCTGCTGCTCAACATGATTCAGAAAACTCAACTAATGT 714
QY 603 ATCTAGAGTACAGTATGATCGGGTTCATATGATACCATCGTCTCGCTCGGATGG 662
Db 715 ATCTAGAGTACAAATGATGTTGGTTCATCTGATACCATGGTCTTCCGCTGGGATGG 774
QY 663 ACTGTTACTATCTCTGATACAAATGCTCTTCAGAAAGTTCATCAGAACGATACATCTCT 722
Db 775 ATTGGTACTATCTCTGATACAACTGCTCTTCAGAAAGTTCATCAGAACGATACATCTCT 834
QY 723 AGCAGTACGAGCTTGGCCATTTACAGTCTCTAGTTGCGAGTGTGTGTCTCATAGG 782
Db 835 AGCAGTCTGGACTTGGCCAAATTACCACTCTAGTTGCTACTTGTGTGTCTCATAGG 894
QY 783 ACTTTTTCAGAGTGGAGTGGGAACTTTGCCAAGTGGATGAGAACTACAACTCGG 842
Db 895 ACTGTTTTCAGAGTGGAGTGGGAACTTTGCCAAGTGGATGAGAACTACAACTCTGG 954
QY 843 GAAAGTGTATATGTTTGTGCTTTAGCTTCGCTCGGAGCTATTTCTGGCAAGTCTACACTCC 902
Db 955 GAAAGTGTATATATCTTGTGCTTTGGCCCTCAGCAGCTATTTCTGGCAAGTATACACTGT 1014
QY 903 TGGTCTTGTGGATGATCTTTCGAGTCACTCTCTGTGTCTTCCAAATTCATAACAGCTGT 962
Db 1015 TGGTGTGTGGATGATTAATCTTTCGAGTCTTCTCTGTGTCTTCCAAATTCATAACAGCTGT 1074
QY 963 GGGATGCTATAGTTCCAGTTCGGGAGTATAGTATTTTCCATGATAGATGGAGCATC 1022
Db 1075 GGGATGCTATAGTTCCAGTTCGGGAGTATAGTATTTTCCATGATAGATGGAGCATC 1134
QY 1023 CAAAATCTTCTCCATTTATTTAGCTATCTCGGGCTTCTTTCATTCTGCTATCAGCACTA 1082
Db 1135 AAAGATTTTCTCCATTTATTTAGCTATCTCGGGCTTCTTTCATTCTGCTATCAGCACTA 1194
QY 1083 CCTCGAGAAAGATTTGAATCTATGAGCCACACAAAGTGTCTGAGAGA 1130
Db 1195 CCTCGAGAAAGATTTGAAGACTTTCGACAGCAAAACCTGTGAGGA 1242
```

```
RESULT 4
CNS0A2S8
LOCUS
DEFINITION
CNS0A2S8 1375 bp mRNA linear HTC 06-FEB-2004
Arabidopsis thaliana Full-length cDNA Complete sequence from clone
GSLTSL202A10 of Adult vegetative tissue of strain col-0 of
Arabidopsis thaliana (thale cress).
ACCESSION
VERSION
BX827187.1 GI:42459879
SOURCE
HTC; GSUT_cDNA.
ORGANISM
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE
1 (bases 1 to 1375)
Castelli,V., Aury,J.M., Jaillon,O., Wincker,P., Clepet,C.,
Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V.,
Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.
Whole Genome Sequences and 'Full-length' cDNA Sequences:
A Combined Approach to Evaluate and Improve Arabidopsis Genome
Annotation
REFERENCE
2 (bases 1 to 1375)
Unpublished
Genoscope.
DIRECT SUBMISSION
Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
- Web : www.genoscope.cns.fr)
COMMENT
The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out
full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli
V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,
Schachter V., Weissenbach J., Salanoubat M.
URGV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis
genome released by MIPS (Munich Information center for Protein
Sequences). 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full
length
http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.
FEATURES
Location/Qualifiers
1..1375
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/strain="Col-0"
/db_xref="taxon:3702"
/clone="GSLTSL202A10"
/tissue_type="Adult vegetative tissue"
/plasmid="pCMVSPORT_6"
1..1375
/gene="At4g18205"
ORIGIN
Query Match 76.1%; Score 909.2; DB 3; Length 1375;
Best Local Similarity 90.2%; Pred. No. 3.5e-249;
Matches 987; Conservative 0; Mismatches 98; Indels 9; Gaps 1;
QY 37 GGTGACCAGAACTTAGAAGCAAAACCTTATAGATCATGAGGTGGTAACTGAATCATCATCA 96
Db 72 GGTGACCAGAACTTAGAAGCAAAACCTTCTAGATCATGAG-----GAACTGAATCA 122
QY 97 TCAGCTGTGCTCAAAACCGAGAACTATAAAAGGTGGTTCGTGTCTCCATATACGTAATC 156
Db 123 TTTTCAGTACCTCAAAACGAAGAACTGTAAAGGTGGCTCGCTGTCTCCATATACGCAATC 182
QY 157 TTTGTCTCTTTTCCAGCCACTAGCTACAAATCTGGGTAGATTGTACTATGAAAATGGA 216
Db 183 TTTGTCTCTTTTCCAGCCAAACCTTGTCTACAGTTCTGGGTAGACTGTACTATGAAATGGA 242
QY 217 GGAATAGCACAATATGTGGTAAACACTTCTTCAACTCATTTGGCTTCCCTGTACTGGTTCTG 276
```

Db	243	GGGAAAGCACATATGTGTGAACACTTCTTCAACTCAITGGGCTTCCTCTACTGATTC	302
Qy	277	TTCCGCTCTTTTCTCGAATCAGGCACCAACCAATCAACAGATACAAATTTCACTCAGTC	336
Db	303	TTCCGCTCTTTTCTCGAATCAGGCACCAACCAATCAACAGATACAAATTTCACTCAGTC	362
Qy	337	CTTCTCTTCCACACCTTCGATCGGTTACTTGTGTGACCTGGACTGTAGTGTCCGCTTAT	396
Db	363	CTTCTCTTCCACACCTTCGATCGGTTACTTGTGTGACCTGGACTGTAGTGTCTGCTTAT	422
Qy	397	GCTTATTTGTGTGACGAGGTTGCTTACTTACCAAGTCTCTACTTTTCTCCCTCATCTTG	456
Db	423	GCTTATTTGTGTGACGAGGTTGCTTACTTACCAAGTCTCTACTTTTCTCCCTCATCTTG	482
Qy	457	GCTCACAGTTGGCTTCACTGCTTTTCTCATATTTCTTAACTCGCAAAAGTTCACT	516
Db	483	GCTCACAGTTGGCTTCACTGCTTTTCTCATATTTCTTAACTCGCAAAAGTTCACT	542
Qy	517	CTTTGATAGTACGTTCTTGTCTTCTCCTACCTGATATCCTCTGCTCTTCTTGTGTCAAC	576
Db	543	CTTTGATAGTACGTTCTTGTCTTCTCCTTACGGTTTCTCTGCTCTCTGCTGTGTCAAC	602
Qy	577	ACTGATTTCAGAAACTCAACTAATGATCTAGAGTACAGTATGATCGGTTTCATATGT	636
Db	603	ACTGATTTCAGAAACTCAACTAATGATCTAGAGTACAGTATGATCGGTTTCATATGT	662
Qy	637	ACCATCGGCTTCCGCTGGGATTTGACTGTACTATCTCTGATACAAATGCTCTTCAGG	696
Db	663	ACCATCGGCTTCCGCTGGGATTTGACTGTACTATCTCTGATACAAATGCTCTTCAGG	722
Qy	697	AAAGTTTTCAGAAAGTACATCTCCTCAGCAGTACGGAATTTGGCAATTTACCAAGTCTCTA	756
Db	723	AAAGTTTTCAGAAAGTACATCTCCTCAGCAGTACGGAATTTGGCAATTTACCAAGTCTCTA	782
Qy	757	GTTGGAGTGTGTAGTCTCATAGACTTTTTCGAAGTGGAGTGGGAATTTTGCCA	816
Db	783	GTTGGAGTGTGTAGTCTCATAGACTTTTTCGAAGTGGAGTGGGAATTTTGCCA	842
Qy	817	AGTGAGATGAGAACTACAAACTCGGAAAGTGTATATGTTTTCGATTTTACCTCGGCA	876
Db	843	AGTGAGATGAGAACTACAAACTCGGAAAGTGTATATGTTTTCGATTTTACCTCGGCA	902
Qy	877	GCTATTTCTCGGAAGTCTACACTCTCTGTCTTGTGGGATGATCTTCGAGTCATCTCT	936
Db	903	GCTATTTCTCGGAAGTCTACACTCTCTGTCTTGTGGGATGATCTTCGAGTCATCTCT	962
Qy	937	GTGTTCTCCAATTCATACAGCTGGGATTCGCTATAGTTCAGTTCGCGAGTGATA	996
Db	963	GTGTTCTCCAATTCATACAGCTGGGATTCGCTATAGTTCAGTTCGCGAGTGATA	1022
Qy	997	GTGTTCTCCAATTCATACAGCTGGGATTCGCTATAGTTCAGTTCGCGAGTGATA	1056
Db	1023	GTGTTCTCCAATTCATACAGCTGGGATTCGCTATAGTTCAGTTCGCGAGTGATA	1082
Qy	1057	TTCTCTTTCATTCGCTATCAGCACTACCTCGACGAAAGAGTTGAATCTAGCCACACA	1116
Db	1083	TTCTCTTTCATTCGCTATCAGCACTACCTCGACGAAAGAGTTGAATCTAGCCACACA	1142
Qy	1117	AGTGTGTAGGAGA	1130
Db	1143	AAACCTGTTGAGGA	1156
RESULT 5			
CNS0A2U8			
LOCUS			
DEFINITION			
Arabidopsis thaliana Full-length cDNA Complete sequence from clone GSLTSL592F07 of Adult vegetative tissue of strain col-0 of Arabidopsis thaliana (thale cress).			
ACCESSION			
BX827470			
VERSION			
BX827470.1 GI:42460504			
KEYWORDS			
HTC; GSLT_cDNA.			
SOURCE			
Arabidopsis thaliana (thale cress)			

ORGANISM	Arabidopsis thaliana		
REFERENCE	Arabidopsis thaliana; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.		
AUTHORS	1 (bases 1 to 1312)		
TITLE	Castelli, V., Aury, J.M., Jaillon, O., Wincker, P., Clepet, C., Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V., Temple, G., Caboche, M., Weissenbach, J., and Salanoubat, M. Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences: A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 1312)		
AUTHORS	Direct Submission		
TITLE	Genoscope.		
JOURNAL	Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqrefgenoscope.cns.fr		
COMMENT	- Web : www.genoscope.cns.fr The sequences are based on single pass reads. Life Technologies (a division of Invitrogen) members carried out full-length libraries construction : Temple G. Genoscope members carried out sequencing and annotation : Castelli V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M. URV INRA : Clepet C., Caboche M. Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich Information center for Protein Sequences). 5 prime and 3 prime are assembled with Phrap. http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full length http://www.genoscope.cns.fr/cgi-bin/ggb?source=Arabidopsis.		
FEATURES	Location/Qualifiers		
source	1..1312		
	/organism="Arabidopsis thaliana"		
	/mol_type="mRNA"		
	/strain="Col-0"		
	/db_xref="taxon:3702"		
	/clone="GSLTSL592F07"		
	/tissue type="Adult vegetative tissue"		
	/plasmid="pCMVSPORT_6"		
gene	1..1312		
	/gene="At4g18205"		
ORIGIN			
Query Match	75.6%	Score 903;	DB 3; Length 1312;
Best Local Similarity	90.4%	Pred. No. 2.1e-247;	
Matches	990;	Conservative 0;	Mismatches 95; Indels 10; Gaps 2;
Qy	37	GGTGACCAAGAACTTAGAAGCAAACTTATAGATCATGAGTGGTAACCTGAATCATCATCA	96
Db	86	GGTGACCAAGAACTTAGAAGCAAACTTATAGATCATGAGTGGTAACCTGAATCATCA	136
Qy	97	TCAGCTGTGCTTCAACCGAGAACTATAAAGGTGGCTTCGTCTCCATATACGTATTC	156
Db	137	TTTTTCAGTACCTCAGACGAAGAAGCTGTAAGAGGTGGCTCCGTGTCTCCATATACGCAATC	196
Qy	157	TTTGTCTCTTTTTCGACCACTAGCTAGTACAAATCTCTGGGTAGATTGTACTATGAAATCGA	216
Db	197	TTTGTCTCTTTTTCGACCACTAGCTAGTACAAATCTCTGGGTAGATTGTACTATGAAATCGA	256
Qy	217	GGAAATAGCACATATGTGTAACACTTTCTTCAACTCATTTGGCTTCCCTGTACTGTTCTG	276
Db	257	GGGAAAGACACATATGTGTAACACTTTCTTCAACTCATTTGGCTTCCCTGTACTGTTCTC	316
Qy	277	TTCCGCTTCTTTTCTCGAATCAGGCAACCCAAATCAACAGATACAAATTTTCAGTCAGTCC	336
Db	317	TTCCGCTTCTTTTCTCGAATCAGGCAACCCAAATCAACAGATACAAATTTTCAGTCAGTCC	376
Qy	337	CTTTCTCTTCCACCCCTTGTGATCGGTTTACTTGTGCACTGGACTGTAGTGTGCGCTTAT	396
Db	377	CTTTCTCTTCCACCCCTTGTGATCGGTTTACTTGTGCACTGGACTGTAGTGTGCGCTTAT	436
Qy	397	GCTTATTGTCTGCAGTAGGGTTGCTCTACTTACCAGTCTCTACTTTCTCCCTCATCTTG	456





```
Db 153 CACTCCCTTCATAGTCAACTCTCTGTTTCTCTTACCGTCTCCTCTCGCTCGCTCGTGGT 212
Qy 573 CAACACTGATTCAGAAAACTCAACTAATGTAATCTAGAGTACAGATGATGATCGGGTTCAAT 632
Db 213 TGACACCGAGTCAACAAGACACAACAATGTCTCTAGACTACAGTATGTGATCGGGTTAT 272
Qy 633 ATGTACCATCGGTGCTTCGGTGGGATGGAGTGTACTATCTCTGTATACAAAATGCTCTT 692
Db 273 ATGCACCATCGGTGCTTCAGCTGGGATAGGAGTCTGCTATCTCTGTATACAACTCCCTCTT 332
Qy 693 CAGGAAGCTTTTCAGAAAGCATATACCTCTCAGCAGTACACGACTTGGCCATTTTACCAGTC 752
Db 333 CAGGAAGCTTTTCAGAGAGCATACATCTCTAGTAGTCAATGACTTGAATCTATCTACCAAGTC 392
Qy 753 TCTAGTTCGAGTGTGTGATTTCTCATAGGACTTTTTCGAAAGTGGAGAGTGGGAACTTT 812
Db 393 TCTAGTTCGAGTGTGTGATTTCTCATTTGAGACTTTTTCGTGAGCGGAGAGTGGAGAACTTT 452
Qy 813 GCCAGTCAGATGAGAAACTCAAACTCGGGAAGTGTCTATGTTTGTGACTTTAGCCTC 872
Db 453 GCCTAGTCAGATGAGAAACTCAAACTTGGGCAAGTGTCTATGTTGTGAGCTTTGGTCTC 512
Qy 873 GGCAGTATTTCTCGGCAAGTCTACACTCTCTGGTCTTGTGGGATTTGATCTTCGAGTCAATC 932
Db 513 TGCAGTATCTCTGGCAAGTCTACACTGTGTGGCTTGGGGGGTGTGAATTTGATCGTC 572
Qy 933 CTCTGTGTTCTCCAAATCCATAACAGCTGTGGGATTTGCTATAGTTCAG--TTGCCGCA 990
Db 573 CTCTGGGCTCTCCAACTCCATAACCGCGTGGGGTGCCTATATAACGGGCCATAAGCT 632
Qy 991 GTGATAGTTTCCATGATAGATGAGCATGCCATCCAAATCTCTCCATTTT 1043
Db 633 GGATAATTTTCCATGAAAAAATGGAAGCGTCCAAAAATCTTCTTCCATCGTTT 685
```

```
RESULT 9
AV828990
LOCUS AV828990 RAF9 Arabidopsis thaliana cdna linear EST 01-APR-2002
DEFINITION mRNA sequence.
ACCESSION AV828990
VERSION AV828990.1 GI:19871050
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 558)
Seki, M., Narusaka, M., Ishida, J., Kamiya, A., Satou, M., Nakajima, M.,
Oono, Y., Sakurai, T., Carninci, P., Kawai, J., Itoh, M., Ishii, Y.,
Arakawa, T., Shibata, K., Shinagawa, A., Muramatsu, M., Hayashizaki, Y.
and Shinozaki, K.
Large scale analysis of Arabidopsis full-length cdna (2002b)
Unpublished (2002)
Contact: Motoaki Seki
Plant Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060
Email: msekic@tc.riken.go.jp
An Arabidopsis full-length cdna library was constructed essentially
as reported previously (Seki et al., 1998). cdna cleaved with BamHI
and XhoI was ligated to modified Lambda FLC-1 vector (Carninci et
al., submitted for publication) digested with BamHI and SalI. This
clone is in a modified pBluescript vector. Please visit our web
site (http://www.gsc.riken.go.jp/e/plant/index\_e.html) for further
details.
```

```
FEATURES
source
Location/Qualifiers
1..558
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
```

```
/db_xref="taxon:3702"
/clone="RAF09-36-N15"
/dev_stage="plants at various developmental stages from
germination to mature seeds"
/lab_host="DH10B"
/clone_lib="RAFL9"
/notes="Site 1: BamHI; Site 2: SalI; subjected to
dehydration (1, 2, 5, 10, 24 hr) and cold (1, 2, 5, 10, 24
hr) treatments"
ORIGIN
Query Match 32.3% Score 386; DB 1; Length 558;
Best Local Similarity 91.0%; Pred. No. 3.5e-99;
Matches 423; Conservative 0; Mismatches 33; Indels 9; Gaps 1;
Qy 37 GGTGACCAGAACTTAGAAGCAAACTTATAGATCATGAGGTGTAACGTGAATCATCATCA 96
Db 103 GGTGACCAGAACTTAGAAGCAAACTTCTAGATCATGAG-----GAAACTGGAATCA 153
Qy 97 TCAGCTGTGCTTCAACCGGAGAACTATAAAGGTGGCTTCGTCTCCATATAGCTAATC 156
Db 154 TTTTCAGTACCTCAACGAGAACTGTAGAGGTGGCTCGCTCTCCATATAGCAATC 213
Qy 157 TTTGTCTCTTTTCCAGCCAGCCACTAGTACAAATCTGGGTAGATTGTACTATGAAAAATGGA 216
Db 214 TTTGTCTCTTTTCCAGCCAGCCACTAGTACAAATCTGGGTAGATTGTACTATGAAAAATGGA 273
Qy 217 GGAATAGCATATGTGGTAAACACTTTCTTCAACTCAATCGGCTTCCCTGACTGGTTCTG 276
Db 274 GGGAAAAGCATATGTGGTAAACACTTTCTTCAACTCAATCGGCTTCCCTGACTGGTTCTC 333
Qy 277 TTCGGCTCTTTTCTCGAATCAGGCAACCCAAATCAACAGATACAAATTTTCAGTCAGTCC 336
Db 334 TTCGGCTCTTTTCTCGAATCAGGCAACCCAAATCAACAGATACAAATTTTCAGTCAGTCC 393
Qy 337 CCTTCCTTCAACCCCTTGCATCGTTTACTTGTGCACTGGACTGTAGTGTCCGCTTAT 396
Db 394 CCTTCCTTCAACCCCTTGCATCGTTTACTTGTGCACTGGACTGTAGTGTCCGCTTAT 453
Qy 397 GCTTATTTGTCTGCAGTAGGGTTGCTCTACTTACCAGTCTCTACTTTTCCCTCAATCTG 456
Db 454 GCTTATTTGTCTGCAGTTGGCTTGTCTTACTTACCANTCTCTACTTTTCTCCTCATCTTG 513
Qy 457 GCCTCAGTGTGGCTTCACTGCTTTTTCATATTTTCTTAAAC 501
Db 514 GCCTCAGTGTGGCTTCACTGCTTTTTCATATTTTCTTAAAC 558
```

```
RESULT 10
AI996743/c
LOCUS AI996743 446 bp mRNA linear EST 08-SEP-1999
DEFINITION 701668184 A. thaliana, Columbia Col-0, root-1 Arabidopsis thaliana
cdna clone 701668184, mRNA sequence.
ACCESSION AI996743
VERSION AI996743.1 GI:5843648
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 446)
```

```
REFERENCE
AUTHORS Chen, J., Miyama, M., Chan, E., Mooney, M., Carroon, B., Gilliland, D.,
Wang, X., Hillman, J., Guegler, K., Kim, C., Doyle, M., Brzoska, P.,
Gorgone, G., Burns, D., Griffin, J., Mouanoutou, M., Nguyen, D.,
Tan, R., Rose, M., Warren, B., Ton, B., Kastury, K., Borillo, C.,
Carpio, T., Policky, J., Suzuki, G., Argentine, C., Shah, S.,
Nobrega, A., Murry, B., Turner, C., Krikorian, S., Elder, L. and
Hanson, D.
Arabidopsis thaliana Gene Expression MicroArray
Unpublished (1999)
Contact: David Smoller, Ph.D.
Genome Systems, Inc., a wholly owned subsidiary of Incyte
```



Pharmaceuticals, Inc.  
4633 World Parkway Circle, St. Louis, MO 63134, USA  
Tel: 877-577-2733  
Fax: 314-427-3324  
Email: service@genomesystems.com.

# FEATURES

source  
1. .446  
Location/Qualifiers  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/ecotype="Col-0"  
/db\_xref="taxon:3702"  
/clone="701668184"  
/tissue\_type="root"  
/dev\_stage="4 - 7 weeks"  
/clone\_lib="A. thaliana, Columbia Col-0, root-1"  
/notes="Vector: pSPORT; Site 1: NotI; Site 2: SalI; cDNA library was derived from untreated root tissue from Arabidopsis thaliana, Columbia Col-0, at 4 - 7 weeks. Plants were grown in 1:1:1 peat moss/vermiculite/perlite soil at 22 deg. C +/- 3 deg. C under constant light, and watered with fertilizer. cDNA synthesis was initiated using a NotI-oligo(dT) primer. Double-stranded cDNA was blunted, ligated to SalI adaptors, digested with NotI, size-selected, and cloned into the NotI and SalI sites of the pSPORT vector."

## ORIGIN

Query Match 28.0%; Score 334.8; DB 1; Length 446;  
Best Local Similarity 99.1%; Pred. No. 1.6e-84;  
Matches 336; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 856 GTTTGACATTAGCCCTCGGCAGCTATTTCCTGCGAAGTCTACACTCTCGTCTGTGGGA 915  
DB 446 GTTTTGACATTAGCCTCGGCAGCTATTTCCTGCGAAGTCTACACTCTGTGGGA 387

QY 916 TTGATCTTCAGTATCCTCTGTGTTCTCAATTCATTAACAGCTGTGGGATTGCTATA 975  
DB 386 TTGATCTTCAGTATCCTCTGTGTTCTCAATTCATTAACAGCTGTGGGATTGCTATA 327

QY 976 GTTCCAGTTGCGGCAGTATAGTTTCCATGATAGATGCGCGCATCCAAATCTTCTCC 1035  
DB 326 GTTCCAGTTGCGGCAGTATAGTTTCCATGATAGATGCGCGCATCCAAATCTTCTCC 267

QY 1036 ATTATTTAGTATCTGCGGCTTCTTTTCATTCGTCTATCAGCACTACCTCGACGAAAG 1095  
DB 266 ATTATTTAGTATCTGCGGCTTCTTTTCATTCGTCTATCAGCACTACCTCGACGAAAG 207

QY 1096 AAGTTGAATACTAGCCACACAGTCTGTAGGAGATCTTCATCTACCTGTGAGGAAGT 1155  
DB 206 AAGTTGAATACTAGCCACACAGTCTGTAGGAGATCTTCATCTACCTGTGAGGAAGT 147

QY 1156 CACACAAACATCAAAAGTGTGTATCAAAAGCATATTTC 1194  
DB 146 CACACAAACATCAAAAGTGTGTATCAAAAGCATATTTC 108

## RESULT 11

CNS0A3B5 1519 bp mRNA linear HTC 06-FEB-2004  
LOCUS Arabidopsis thaliana Full-length cDNA Complete sequence from clone GSLTFB632D12 of flowers and buds of strain col-0 of Arabidopsis thaliana (thale cress).  
BX826795  
BX826795.1 GI:42460970  
HTC; GSLT cDNA.  
DEFINITION Arabidopsis thaliana (thale cress)  
SOURCE Arabidopsis thaliana  
ORGANISM Arabidopsis thaliana  
KEYWORDS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
1 (bases 1 to 1519)  
REFERENCE Castelli, V., Aury, J.M., Jaillon, O., Wincker, P., Clepet, C., Menard, M., Cruaud, C., Quetier, P., Scarpelli, C., Schachter, V.,

## TITLE

JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL

## COMMENT

Temple, G., Caboche, M., Weissenbach, J. and Salanoubat, M.  
Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences:  
A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation  
Unpublished  
2 (bases 1 to 1519)  
Genoscope.  
Direct Submission  
Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)  
The sequences are based on single pass reads.  
Life Technologies (a division of Invitrogen) members carried out full-length libraries construction : Temple G.  
Genoscope members carried out sequencing and annotation : Castelli V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach, J., Salanoubat M.  
URV INRA : Clepet C., Caboche M.  
Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich Information center for Protein Sequences). 5 prime and 3 prime are assembled with Phrap.  
http://www.genoscope.cns.fr/externe/sequences/Banque\_Projet\_EF/Full\_length  
http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.

## FEATURES

source

Location/Qualifiers  
1. .1519  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/strain="Col-0"  
/db\_xref="taxon:3702"  
/clone="GSLTFB632D12"  
/tissue\_type="Flowers and buds"  
/plasmid="pCMVSPORT\_6"  
1. .1519  
/gene="At4g18210"

## ORIGIN

Query Match 26.5%; Score 317; DB 3; Length 1519;  
Best Local Similarity 59.7%; Pred. No. 3.1e-79;  
Matches 641; Conservative 0; Mismatches 420; Indels 12; Gaps 6;

QY 96 ATCAGCTGTGCTCAAAACGAGAACTATAAAGGTGGCTTCGTGTCATATACGTA-A 154  
DB 176 AGCAAGATATCTTACTCTAAACATACAAACGGTGGCTCAGGGTGACTTATATACATT 235

QY 155 TCTTTGTCCTCTTTTGGCCAGCCACTAGTACAAATCTCGGTAGATTGTACTATGAATG 214  
DB 236 TCTTTGTCAATTCAGGCCAAACAGTTGCTACAAATTTGGCGAGAGTACTATGACAAG 295

QY 215 GAGAAATAGCACAT-ATGTGGTAACACTTCTTCAACTCATTTGGCTTCCCTGTACTGGTT 273  
DB 296 GAGGAACAGTAATGGCTTAGCAACGGTAGTTCAACTTTGTGGCTTCTCTGTCTACTT 355

QY 274 CTGTTCCGCTCTTTTCTCGAATCAGGCAACCCAAATCAACAGATACAAATTTTCACTCAG 333  
DB 356 CCATATTATATCTTGTCTATTTA-----AAACACATGCAACAACTGATAGATGAAAA 409

QY 334 TCCCTTCTTCAACACCTTGCATCGGTTTACTTGTGCACCTGGACTGCTAGTGTCCGCT 393  
DB 410 AGAACCTCACCCTAGGAACCGTATTTGGTTTACGTAAGTGTGGACTTCTTGTAGGAGCA 469

QY 394 TATGCTTATTTGTCTGCAAGTGGGTTGCTCTACTTTACCAGTCTCTACTTTTCTCCCTCATC 453  
DB 470 GATTGCTATTTGTACTCCATTTGGACTTCTTACTTACCGGTTTCTACCTATTTCCCTGATC 529

QY 454 TTGGCCTCAGAGTTGGCCTTCACCTGCCTTTTTCATATTTTCTTAACTCGCAAAAGTTC 513  
DB 530 TGTGCTCTCAGTTAGCCTTCAATGCTTTTCTTCTTATTTTCTTAACTCACAAAAACTT 589

QY 514 ACTGCTTTGATAGTCAAGTTCTTTGCTTCTCTCTCACTGTATCTCTGTCTTCTTGTGTC 573  
DB 590 ACCCTATCATTTTAAATTTCTTTTCTCTCTCACTATATTTTCCACCTCTACTTGCATTC 649

```
Qy 574 AACACTGATTCAGAAACCACTAATATGTATCTAGAGTACAGTATGTGATCGGGTTTCATA 633
|||
Db 650 AATATGAGGAGACAGATCCACAAAGTTACAAAGGAGATTTTTTAAAGGTTTCATA 709
|||
Qy 634 TGTACATCGGTGCTCCGCTGGGATGAGCTGTTACTATCTGATACA-AAATGCTCTT 692
|||
Db 710 TGCACCGTTGCTGCTGTTGCTGGTTATGTTTATTTATCCCTACAAACGAGGCTTT 769
|||
Qy 693 CAGGAAAGTTTTCAGAAAGCATACATCTCAGCAGTACGAGCTTGGCCATTTACCAGTC 752
|||
Db 770 TTTAAAGTCTTAAAGAGCAAAATTTCTCAGAAGTTATGGATATGATAATCTACATGAA 829
|||
Qy 753 TCTAGTTCGAGTGTGTGATTTCTCATAGGACTTTTTGCAAGTGGAGAGTGGGAAACTTT 812
|||
Db 830 TCTGATGGGAGTTGTTGCGGTGGTGGGCTTTTGTCTCACAATGAATGGAATAATTT 889
|||
Qy 813 GCCAAGTGCAGATAGAGAACTCAAACTCGGAAAGTGTCTATGTTTGTGACTTTAGGCTC 872
|||
Db 890 GAACAATGAAAGGAAACTAAAAAGGAAAGTAACCTACAATATGAACCTAAATGTG 949
|||
Qy 873 GGCAGCTATTTCCCTGGCAAG--TCTACACTCTGTCTGTGGGATTCATCTCGAGTCA 930
|||
Db 950 AACAGCTGTACCTGGCCGCTATCTCCAGTGGGAAACAAGAACTGAATCTTCGAGCTC 1009
|||
Qy 931 TCCTCTGTGTTCTCCAAATTCAT-AACAGCTGTGGGATTCCTATAGTTCCAGTTGCGGC 989
|||
Db 1010 TCCCTCTATTTTAAATGCAATAAAGGCTTTTGGAACTCCAGTGGTCTCTATCTTGGC 1069
|||
Qy 990 AGTGATAGTTTCCATGATAGAAAGGACGCATCCAAATCTTCTCCATATTTTATGCTAT 1049
|||
Db 1070 TGTATCATTTTCCATGACAAATAAATGGGTAAAGGTGATTTCTATGATCTCTAGCTAT 1129
|||
Qy 1050 CTGGCGCTTCCTTCATCTGCTATCAGCAGTCTCTCGAACGAAAGAGTGAATACTAG 1109
|||
Db 1130 TTGGGGTTTAACTTCCATGCTACCAAAAATATCTTGATCAGAAAAACTTGAAGAAAA 1189
|||
Qy 1110 CCACAAAGTGTGTAGGAGATCTTCATCTACCTGTTGAGGAAGTCCACAA 1162
|||
Db 1190 TCATGAATATACAAACAGAAATCCCTGACCCAGCAAGCAGAGAGTCA 1242
|||

RESULT 12
CNS0A36P 1323 bp mRNA linear HTC 06-FEB-2004
LOCUS Arabidopsis thaliana Full-length cDNA Complete sequence from clone
DEFINITION GSTTFB64ZF05 of Flowers and buds of strain col-0 of Arabidopsis
thaliana (thale cress) .
ACCESSION BX826811
VERSION BX826811.1 GI:42460388
KEYWORDS HTC; GSLT cDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1 (bases 1 to 1323)
Castelli,V., Aury,J.M., Jaillon,O., Wincker,P., Clepet,C.,
Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V.,
Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.
Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences:
A Combined Approach to Evaluate and Improve Arabidopsis Genome
Annotation
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1323)
GENOSCOPE Genoscope.
TITLE Direct Submission
JOURNAL Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqrefegenoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT The sequences are based on single pass reads.
Life technologies (a division of invitrogen) members carried out
full-length librairies construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli
```

```
V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,
Schachter V., Weissenbach J., Salanoubat M.
URGV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis
genome released by MIPS (Munich Information center for Protein
Sequences). 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/Sequences/Banque_Projet_EF/Full
length
http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.
location/Qualifiers
1. 1323
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/strain="Col-0"
/db_xref="taxon:3702"
/clone="GSLTFB64ZF05"
/tissue_type="Flowers and buds"
/plasmid="pCMVSPORT_6"
1. 1323
/gene="At4g18190"
ORIGIN
Query Match 26.2%; Score 312.8; DB 3; Length 1323;
Best Local Similarity 56.5%; Pred. No. 4.7e-78;
Matches 609; Conservative 0; Mismatches 457; Indels 12; Gaps 1;
Qy 53 AAGCAACCTTATAGATCATGAGTGTGTAAGTGAATCATCATCATGAGTGTGCCTCAA 112
|||
Db 27 AATCAAGACTCCACTCAGTAAATGGCGACTGAAGGGAATTTTCAACAGAGAGAGAA 86
|||
Qy 113 CCAGAAACTATAAAGGTGGCTTCGTCTCTCCATATACGTAATCTTTGTCTCTTTTGGC 172
|||
Db 87 GTCAAGATCTCTTGAGGTTAAGAGTGTCTCTATGTCACCTCTCTCTAGCTGAG 146
|||
Qy 173 AGCCACTAGTACAATTTCTGGGTAGATTGTATATGAAATGAGAGAAATGACATATG 232
|||
Db 147 AGCAATAGGCACTCTCTTAGGTAGACTTTTACTACGAAAAAGCGGTAAAGACATGCG 206
|||
Qy 233 TGGTAACACTTCTCAACTCATTTGGCTTCCTCTGCTACTGTTCTGTCGGTCTCTTTCTC 292
|||
Db 207 TCGAAACCTTGGTTACGTTGTAGGGTTTCCTTTAACCTTCTCTGCTATTATTACTTAA 266
|||
Qy 293 GAATCAGGCAACCAATCAACAGATACAAATTTCACTAGTCCCTTCTCTCACCACCC 352
|||
Db 267 AGCTGAGCGCTCCAGACTTAAACCACTTACCAAAAAAATCTCTCTCTCTTGACAC 326
|||
Qy 353 TTGCATCGGTTTACTTTGTGCATGAGTGTCTGCTTATGCTTATGTTTGTCTGAG 412
|||
Db 327 TATCTTTAGTGATATTGGACTTGGCTTGGCTTGGCTGGACATTTGATTTTGTACTCAT 386
|||
Qy 413 TAGGGTTGCTCTACTTACCACTGCTCTACTTTCTCCCTCATCTTGGCCTCACAGTTGGCCT 472
|||
Db 387 TTGGGCTACTTTACCTTCTCTCAACTTTCTTTTGATCTCTGCGTCGCAATTTGGCTT 446
|||
Qy 473 TCACTGCTTTTCTCATATTTCTTAACTCGCAAAAGTTCACTCTCTTGATAGTCACTT 532
|||
Db 447 TTAAGCGGCTCTCTCTTACTTCTTAACCTACAAAAAATCAACCACTTTATATCAATT 506
|||
Qy 533 CTTTGTCTTCTCCTACGTATCTCTGCTCTCTCTTGTGGTCAACACTGATTCAGAA---- 589
|||
Db 507 CACTTGTCTCTTAAACCATATCTTACACTTCTTGTATTATCCAAACATGAACCAAACTC 566
|||
Qy 589 -----NACTCAACTAATGTATCTAGTACAGTATGTATGATCGGGTTCATATGTACCA 640
|||
Db 567 CTTCTTCTCTTCAAGTCCGAGCCAGTCCAAAGTATGATTAGATATCATCTCGCCGG 626
|||
Qy 641 TCGGTGCTTCGGCTGGGATTTGGACTGTTTACTATCTCTGTATACAAATGCTCTTCAGGAAAG 700
|||
Db 627 TCGGTAGCTCAGCTGGTTATTTCTCTGGTGTCTTTTAAACAGATTACGCGTACGAAAGA 686
|||
Qy 701 TTTTCAGAAAGCATACATCTCTAGCAGTACGAGACTTGGCACTTACCACTCTCTAGTTG 760
|||
Db 687 TTCAAAAGAAATACACATTTCAAGGCTATTTTAGACATGGCCACATATCTCCGCTCTATGGTAG 746
|||
```

QY 761 CGAGTTGTGTAGTTCTCATAGGACTTTTTCAGAGTGGAGTGGGAAACTTTGCCAAGTG 820  
 |||||  
 Db 747 CCACCTTGTGTAGTTGTGGTAGGACTTTTGGAGTGGTGGTGGGAAAGCTGAGTACAG 806  
 |||||  
 QY 821 AGATGAGAACTACAACTCGGAAAGTGCATATGTTTGTGACTTTTAGCTCGGAGCTA 880  
 |||||  
 Db 807 AATGGAAGAGTTTCAACTAGGGAAGCTCATACATTTTGTAAACATCGGTTCACGA 866  
 |||||  
 QY 881 TTTCCTGGCAAGTCTACACTCTCGGTCTTGTGGGATTGATCTTCGAGTCACTCTGTGT 940  
 |||||  
 Db 867 TATCATGGCAAGCTTGTGATTGGAAGTGTGTGTTGATTATCGAAGTTTTCATCGCTTT 926  
 |||||  
 QY 941 TCTCCAATTCCATAACAGCTGTGGGATTCCTATAGTTCAGTTGGCGGAGTGATAGTTT 1000  
 |||||  
 Db 927 TTTCCAATATCATAGAACTCTTTGTTTACAAAGTTGCTGCTGTTCTTGTGTTATCTTCT 986  
 |||||  
 QY 1001 TCCATGATAGATGGAGCATCCAAATCTTCTCCATTATTTAGTATCTCGGGCTTCC 1060  
 |||||  
 Db 987 TCGTGATAGATAAGTGAATCAAGTTGGTGGCAATGTTTGGCCATCTCGGGATTG 1046  
 |||||  
 QY 1061 TTTCATTCCTCTATCAGCACTACTCGACGAAAGAGTTGAATCTAGCCACACAAG 1118  
 |||||  
 Db 1047 TTTCATTGTTATCAGCATTTATGTCATGATAGAAAGCAAGAACCAAGAG 1104  
 |||||

RESULT 13  
 AU226302/c 498 bp mRNA linear EST 23-APR-2002  
 LOCUS AU226302 RAFL14 Arabidopsis thaliana cDNA clone RAFL14-23-A22 3',  
 DEFINITION mRNA sequence.  
 ACCESSION AU226302  
 VERSION AU226302.1 GI:19740949  
 KEYWORDS EST.  
 SOURCE Arabidopsis thaliana (thale cress)  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 1 (bases 1 to 498)  
 Akiyama,K., Enju,A., Oono,Y., Sakurai,T., Carninci,P., Kawai,J.,  
 Itoh,M., Ishii,Y., Arakawa,T., Shibata,K., Shinagawa,A.,  
 Muramatsu,M., Hayashizaki,Y. and Shinozaki,K.  
 Large scale analysis of Arabidopsis full-length cDNA  
 Unpublished (2002)  
 TITLE Contact: Motoaki Seki  
 JOURNAL Plant Functional Genomics Research Group  
 COMMENT RIKEN Genomic Sciences Center  
 3-1-1 Koyadai, Teikuba, Ibaraki 305-0074, Japan  
 Tel: 81-298-36-4359  
 Fax: 81-298-36-9060  
 Email: mseki@rtc.riken.go.jp  
 An Arabidopsis full-length cDNA library was constructed essentially  
 as reported previously (Seki et al., 1998). cDNA cleaved with BamHI  
 and XhoI was ligated to modified Lambda FLC-1 vector (Carninci et  
 al., submitted for publication) digested with BamHI and SalI. This  
 clone is in a modified pBluescript vector. Please visit our web  
 site ([http://www.gsc.riken.go.jp/e/plant/index\\_e.html](http://www.gsc.riken.go.jp/e/plant/index_e.html)) for further  
 details.

FEATURES  
 source  
 1. .498  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3702"  
 /clone="RAFL14-23-A22"  
 /tissue\_type="root"  
 /lab\_host="DH10B"  
 /clone\_lib="RAFL14"  
 /note="Site\_1: BamHI; Site\_2: SalI"

ORIGIN  
 Query Match 25.1%; Score 299.6; DB 1; Length 498;  
 |||||

Best Local Similarity 98.7%; Pred. No. 2.1e-74;  
 Matches 302; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
 QY 889 CAAGTCTACACTCCTGGTCTTGTGGGATTGATCTTCGAGTCACTCTCTGTTCTCCCAAT 948  
 |||||  
 Db 498 CAGGCTTACACTCTTGGTCTTGTGGGATTGATCTTCGAGTCACTCTCTGTTCTCCCAAT 439  
 |||||  
 QY 949 TCCATAACAGCTGTGGGATTGCCTATAGTTCAGATTGGCGGAGTATAGTTCATGAT 1008  
 |||||  
 Db 438 TCCATAACAGCTGTGGGATTGCCTATAGTTCAGATTGGCGGAGTATAGTTCATGAT 379  
 |||||  
 QY 1009 AGAATGGAGCGATCCAAATCTTCTCCATTATTTAGCTATCTCGGGCTTCCTTTTCATTC 1068  
 |||||  
 Db 378 AGAATGGAGCGATCCAAATCTTCTCCATTATTTAGCTATCTCGGGCTTCCTTTTCATTC 319  
 |||||  
 QY 1069 GTCTATCAGCACTACTCTCGAGAAAGAGTTGAATCTAGCCACACAAAGTCTGTAGGA 1128  
 |||||  
 Db 318 GTCTATCAGCACTACTCTCGAGAAAGAGTTGAATCTAGCCACACAAAGTCTGTAGGA 259  
 |||||  
 QY 1129 GATCTTCATCTACCTGTTGAGGAAGTCCACACAAATGTTGTGATCAAAAGCAT 1188  
 |||||  
 Db 258 GATCTTCATCTACCTGTTGAGGAAGTCCACACAAATGTTGTGATCAAAAGCAT 199  
 |||||  
 QY 1189 ATTTCC 1194  
 |||||  
 Db 198 ATTTCC 193  
 |||||

RESULT 14  
 BZ967767  
 LOCUS BZ967767  
 DEFINITION PUDGUS5TD ZM\_0.6.1.0 KB Zea mays genomic clone ZMMBTa184115,  
 genomic survey sequence.  
 ACCESSION BZ967767.1 GI:29185853  
 VERSION BZ967767  
 KEYWORDS GSS.  
 SOURCE Zea mays  
 ORGANISM Zea mays  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
 clade; Panicoideae; Andropogoneae; Zea.  
 1 (bases 1 to 922)  
 Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,  
 Renwick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and  
 Bennetzen,J.  
 Maize Genomics Consortium  
 Unpublished (2003)  
 TITLE Contact: Cathy Whitelaw  
 JOURNAL TIGR  
 COMMENT 9712 Medical Center Drive, Rockville, MD 20850, USA  
 Tel: 301-838-5843  
 Fax: 301-838-0208  
 Email: whitelaw@tigr.org  
 Seq primer: 1P  
 Class: sheared ends.  
 Location/Qualifiers  
 1. .922  
 /organism="Zea mays"  
 /mol\_type="genomic DNA"  
 /strain="B73"  
 /db\_xref="taxon:4577"  
 /clone="ZMMBTa184115"  
 /clone\_lib="ZM 0.6.1.0 KB"  
 /note="Vector: pCR4-TOPO; Site\_1: EcoRI; 0.6-1.0 kb high  
 Cot selected genomic DNA library"

FEATURES  
 source  
 1. .922  
 /organism="Zea mays"  
 /mol\_type="genomic DNA"  
 /strain="B73"  
 /db\_xref="taxon:4577"  
 /clone="ZMMBTa184115"  
 /clone\_lib="ZM 0.6.1.0 KB"  
 /note="Vector: pCR4-TOPO; Site\_1: EcoRI; 0.6-1.0 kb high  
 Cot selected genomic DNA library"

ORIGIN  
 Query Match 20.1%; Score 239.8; DB 8; Length 922;  
 Best Local Similarity 58.2%; Pred. No. 3.7e-57;  
 Matches 421; Conservative 0; Mismatches 302; Indels 0; Gaps 0;  
 QY 376 CGACTGTAGTGTGCGCTTATGCTATTGTTCTGCAGTAGGGTCTCTACTACAGTTC 435  
 |||||

